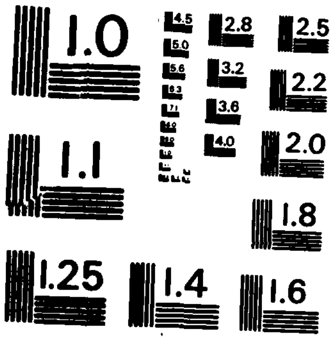
AD-A166 477 POWER OF STATISTICAL TESTS USED IN CORRELATION TECHNIQUES FOR BATTLEFIELD IDENTIFICATION(U) JET PROPULSION LAB PASADENA CA J MYHRE ET AL. AUG 85 1/2

UNCLASSIFIED JPL-D-2793 NAS7-918 F/G 12/1 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

JPL D-2793

INSTITUTE OF DECISION SCIENCE
FOR BUSINESS & PUBLIC POLICY

POWER OF STATISTICAL TESTS USED IN CORRELATION TECHNIQUES FOR BATTLEFIELD IDENTIFICATION

Conducted At

Institute of Decision Science
Claremont McKenna College

And

Mathematical Analysis Research Corporation

For

Jet Propulsion Laboratory

August 1985

Claremont McKenna College
Claremont, California 91711

UAA-016

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER JPL D-2793	2. GOVT ACCESSION NO. AD-A166 477	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) POWER OF STATISTICAL TESTS USED IN CORRELATION TECHNIQUES FOR BATTLEFIELD IDENTIFICATION		5. TYPE OF REPORT & PERIOD COVERED Final 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Dr. Janet Myhre, Michael Rennie, Will Duquette, Daniel Hockman		8. CONTRACT OR GRANT NUMBER(s) NAS7-918
9. PERFORMING ORGANIZATION NAME AND ADDRESS Institute of Decision Science Claremont McKenna College Claremont, CA 91711 (Cont.)		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS RE 182 AMEND # 187
11. CONTROLLING OFFICE NAME AND ADDRESS Commander, USAICS ATTN: ATSI-CD-SF Ft. Huachuca, AZ 85613-7000		12. REPORT DATE August, 1985 13. NUMBER OF PAGES 127
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Jet Propulsion Laboratory California Institute of Technology 4800 Oak Grove Drive, ATTN: 126-200 Pasadena CA 91109		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Dissemination

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Prepared under contract to JPL for the U. S. Army Intelligence Center and School's Combat Developer's Support Facility

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

STATISTICAL TESTS,STATISTICAL DISTRIBUTIONS, POWER OF TESTS, ELLIPTICAL ERROR PROBABLE (EEP), CONFIDENCE ELLIPSES, COVARIANCE MATRICES, MATHEMATICAL SIMULATIONS, EMITTER CORRELATION, ROBUSTNESS, VARIANCE ASSUMPTIONS, SAMPLE SIZES

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This is one of a series of algorithm analysis reports performed for the US Army Intelligence Center and School covering selected algorithms in existing or planned Intelligence and Electronic Warfare (IEW) systems. This report documents an empirical examination of the robustness of several statistical tests. The behaviour of the tests under varying degrees of failure of their underlying assumptions is documented and analyzed.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(Continuation of block 9)
Mathematical Analysis Research Corporation
4239 Via Padova
Claremont CA 91711

(1)

U.S. ARMY INTELLIGENCE CENTER AND SCHOOL
Software Analysis and Management System

Power of Statistical Tests Used in Correlation Techniques for Battlefield Identification

July 1985

by

Dr. Janet Myhre
Mr. Michael Rennie
Mr. Will Duquette
Mr. Daniel Hockman

Concur:

James W. Gillis, Technical Staff
USAMS Task Team

Edward J. Records, Manager
USAMS Task Team

J. P. McClure, Manager
Ground Data Systems Section

Yukio Nakamura, Manager
Defense Information Systems Program

JET PROPULSION LABORATORY
California Institute of Technology
Pasadena, California

86 4 8 - 068

PREFACE

The work described in this publication was sponsored by the United States Army Intelligence Center and School. The writing and publication of this paper was supported by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

POWER OF STATISTICAL TESTS USED IN CORRELATION TECHNIQUES FOR BATTLEFIELD IDENTIFICATION

CONDUCTED AT

INSTITUTE OF DECISION SCIENCE
CLAREMONT MCKENNA COLLEGE

AND

MATHEMATICAL ANALYSIS RESEARCH CORPORATION

FOR

JET PROPULSION LABORATORY

AUGUST 1985

PREPARED BY

Professor Janet Myhre
Will Duquette

TABLE OF CONTENTS

(-a)

ABSTRACT

This report is the second in a series presenting a study of self-correlation algorithms in intelligence systems. It was performed by the Institute of Decision Science at Claremont McKenna College in support of the Algorithm Analysis subtask of the U.S. Army Intelligence Center and School (USAICS), Software Analysis and Management System (USAMS) task at Jet Propulsion Laboratory. These self-correlation algorithms use multivariate statistical tests to determine the equality of mean vectors from two different data sets. For example, tests are used to determine the equality of location vectors from two different data sets (Are the data from the same emitter?). This report considers estimation of the probability that these tests may lead to an incorrect decision. Possible test errors are studied under different assumptions concerning:

1) the distribution of the data, e.g. normal error, skewed error, etc.

2) the estimated location of the emitter, e.g. mean of the data, most frequent value, etc.

3) the variability of the error, i.e. the variance-covariance of the data

4) the amount of data, i.e. sample size

Frequency of test errors were estimated by simulation for most of the cases studied. The results indicate that in some of the cases the error rate is high enough to be of possible concern.

Keywords: statistical distributions

I. INTRODUCTION

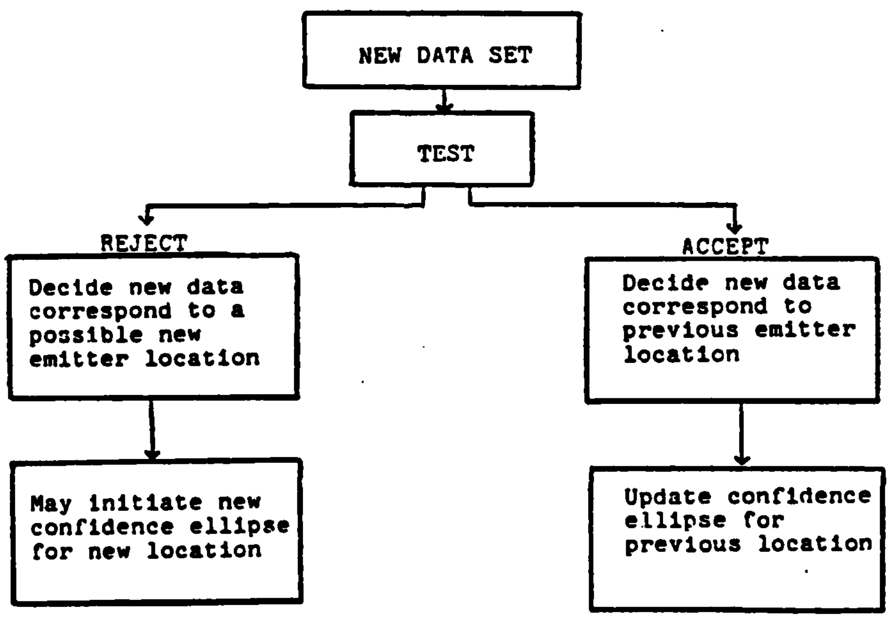
Correlation techniques for battlefield identification use various tests for determining whether or not a new set of data is from a previously identified emitter location. These decision tests are based on certain assumptions about the data. Even if all the assumptions are met the tests have possible errors associated with them. When some of the assumptions concerning the data are not met the frequency of test errors may become even greater, perhaps unacceptable.

This report considers

a. Types of decision errors possible when performing statistical tests.

b. Sensitivity of the frequency of test errors to incoming data which do not meet the assumptions of the statistical tests.

TYPES OF DECISION ERRORS POSSIBLE WHEN PERFORMING STATISTICAL TESTS



One of the inherent problems with statistical tests is that one cannot claim to have made the correct decision with complete certainty. Two types of decision errors can be made:

I. Deciding that the new data <u>are not</u> from the previously identified location when in fact they <u>are</u>.

II. Deciding that the new data <u>are</u> from the previously identified location when in fact they <u>are not</u>.

		TRUE STATE OF NATURE: Same Location	TRUE STATE OF NATURE: Different Location
TEST DECISION	Same Location	No Error	Error of the Second Type
TEST DECISION	Different Location	Error of the First Type	No Error

The statistical tests considered in this report are constructed so that the probability of committing an error of Type I is set at some level, say 0.05.* If we commit an error of the first type we would be putting a new emitter location erroneously into the data-base. If we commit an error of the second type (Type II) we would be erroneously shrinking the confidence ellipse which is centered at the estimated location and also would not be putting the new emitter into the data base.**

The general diagram on page 3 illustrates errors in relation to emitter locations and confidence ellipses. The graph in Figure 1 shows the relationship between a variance-covariance matrix (sensor error) and the resulting 95% confidence ellipse. Figure 2 contains specific examples which describe statistical decisions to relation to assumed emitter locations and confidence ellipses. Additional graphs may be found in the body of the report.

---

* Often erroneously called a test at the 95% level.

** It could also be possible that the new data belong to a different previously identified emitter. This error will not be illustrated here.

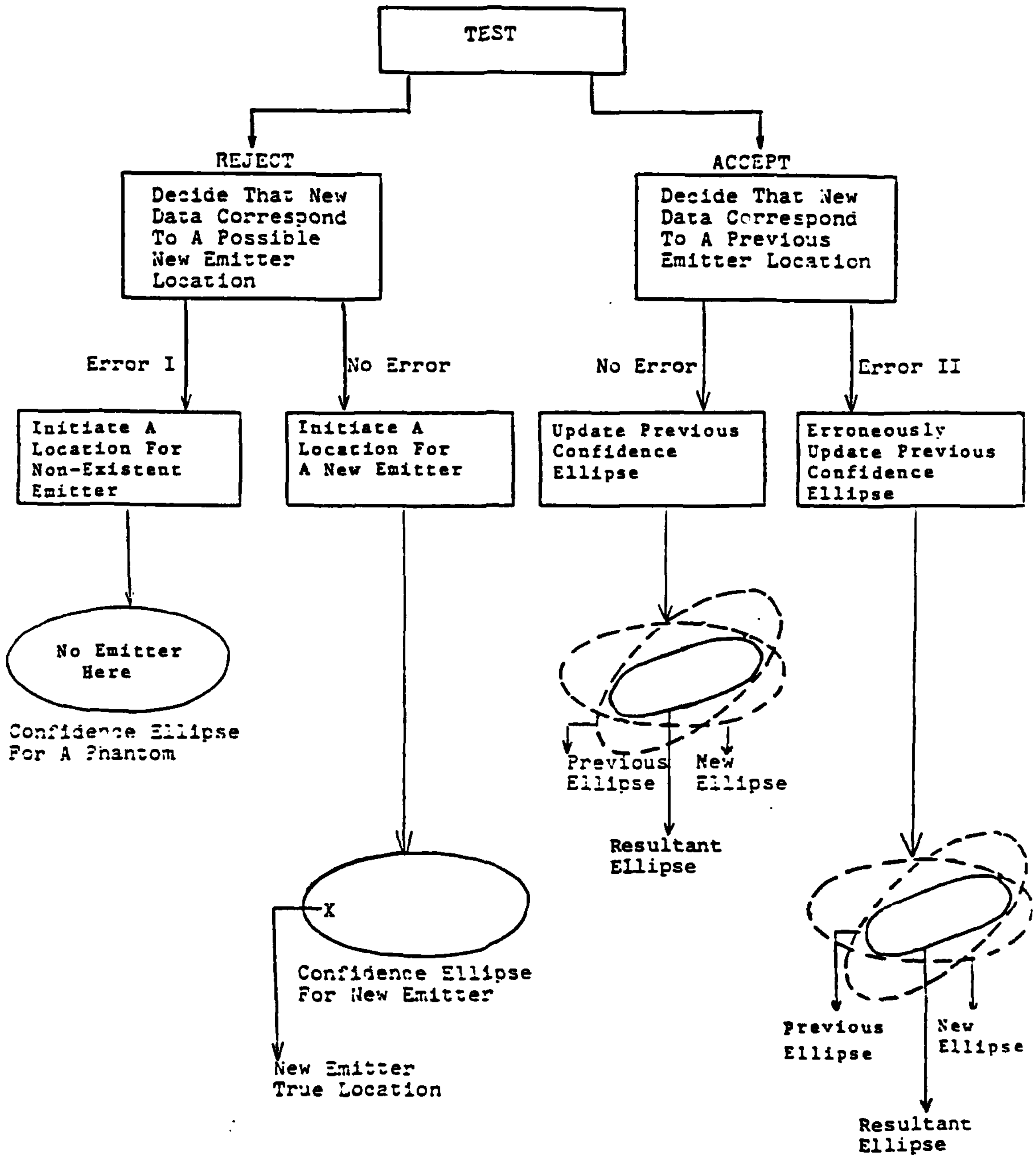
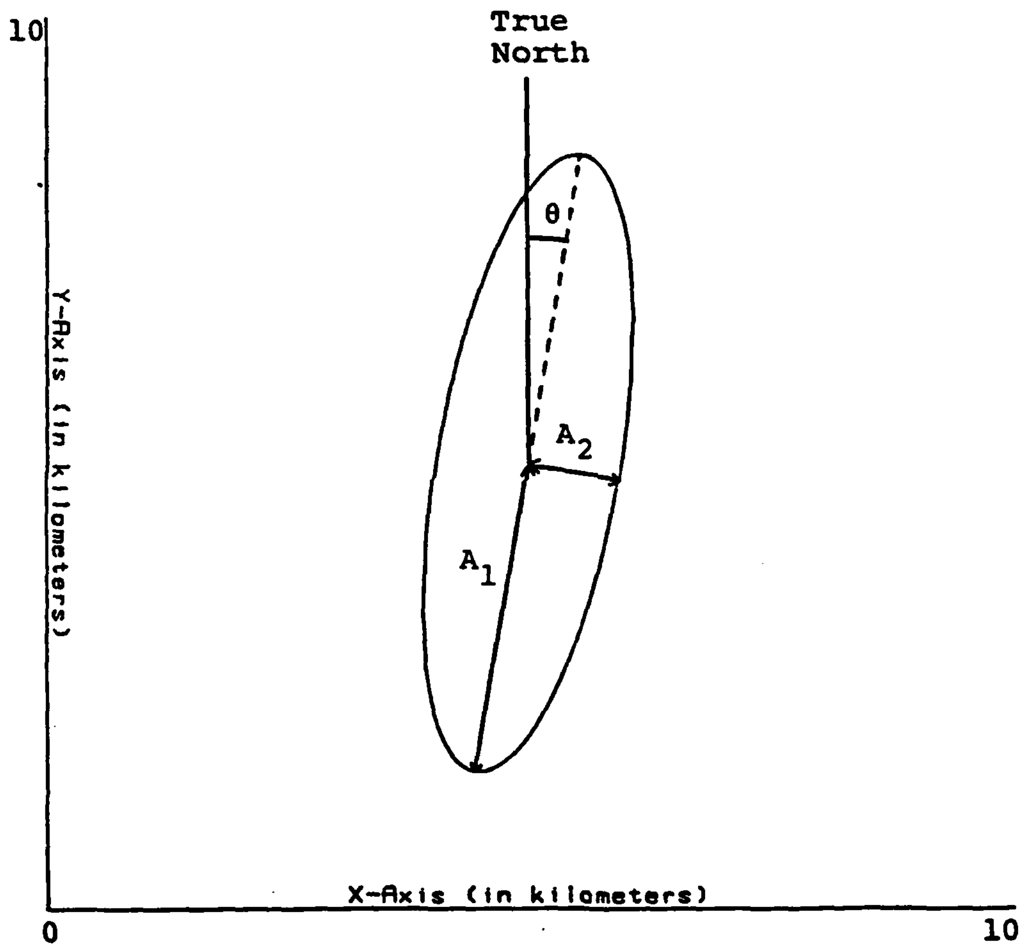

TEST
REJECT
Decide That New Data Correspond To A Possible New Emitter Location
ACCEPT
Decide That New Data Correspond To A Previous Emitter Location
Error I
No Error
No Error
Error II
Initiate A Location For Non-Existent Emitter
Initiate A Location For A New Emitter
Update Previous Confidence Ellipse
Erroneously Update Previous Confidence Ellipse
No Emitter Here
Confidence Ellipse For A Phantom
Previous Ellipse
New Ellipse
Resultant Ellipse
X
Confidence Ellipse For New Emitter
New Emitter True Location
Previous Ellipse
New Ellipse
Resultant Ellipse

FIGURE 1

The spread of the measurement error, shown by the Error Variance Matrix below, determines the shape of the confidence ellipse. Size is determined by the confidence level. The ellipse below is a 95% confidence ellipse.



Error Variance

$$\sum = \begin{bmatrix} 0.2 & 0.3 \\ 0.3 & 2.0 \end{bmatrix}$$

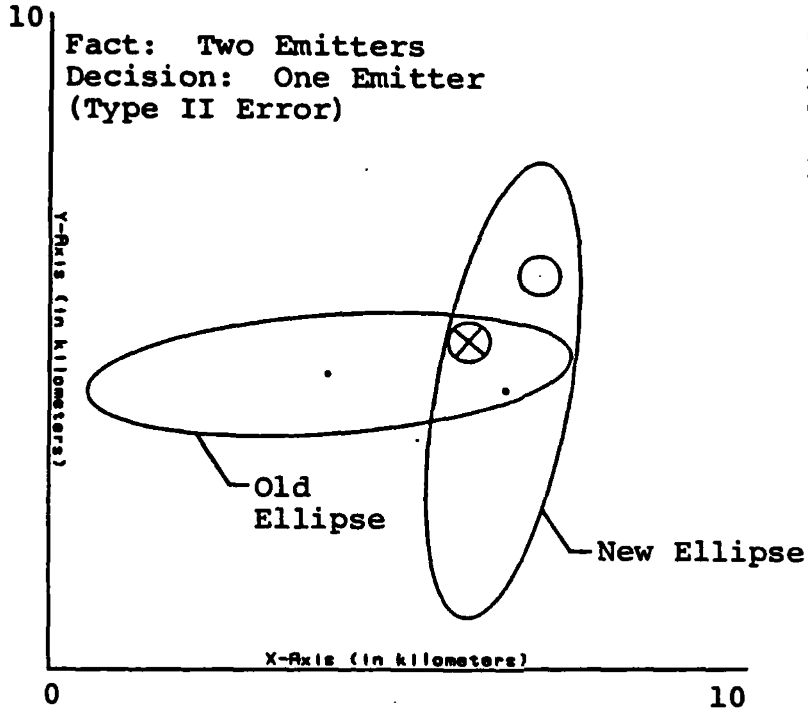
Axes and Orientation

$A_1 = 3.503$

$A_2 = 0.952$

$\theta = 9.2175°$

FIGURE 2

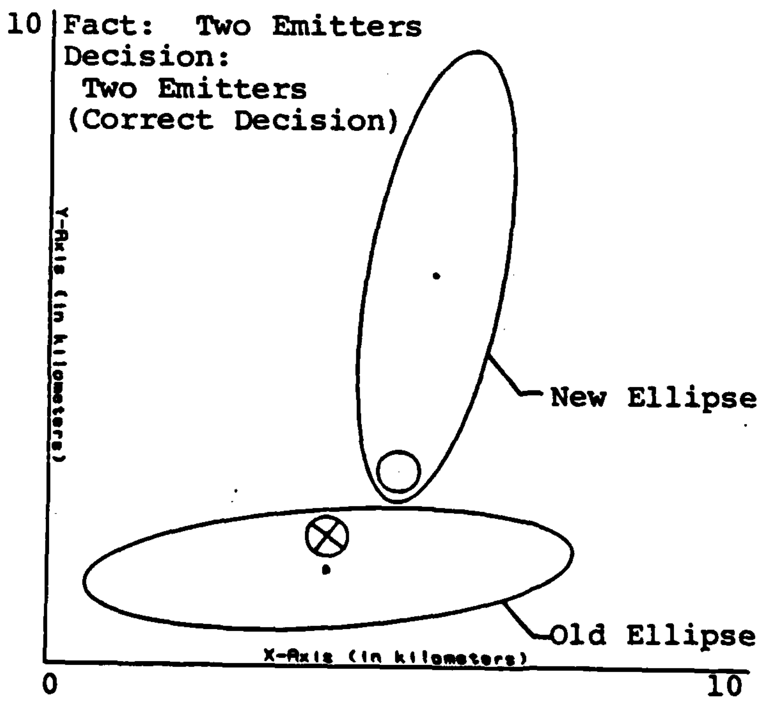


The Database contains position estimates and ellipses; the exact location of the original Emitter is not known.

Error Variances

$$\sum_{\text{Old}} = \begin{bmatrix} 2.00 & 0.15 \\ 0.15 & 0.15 \end{bmatrix}$$

$$\sum_{\text{New}} = \begin{bmatrix} 0.2 & 0.3 \\ 0.3 & 2.0 \end{bmatrix}$$

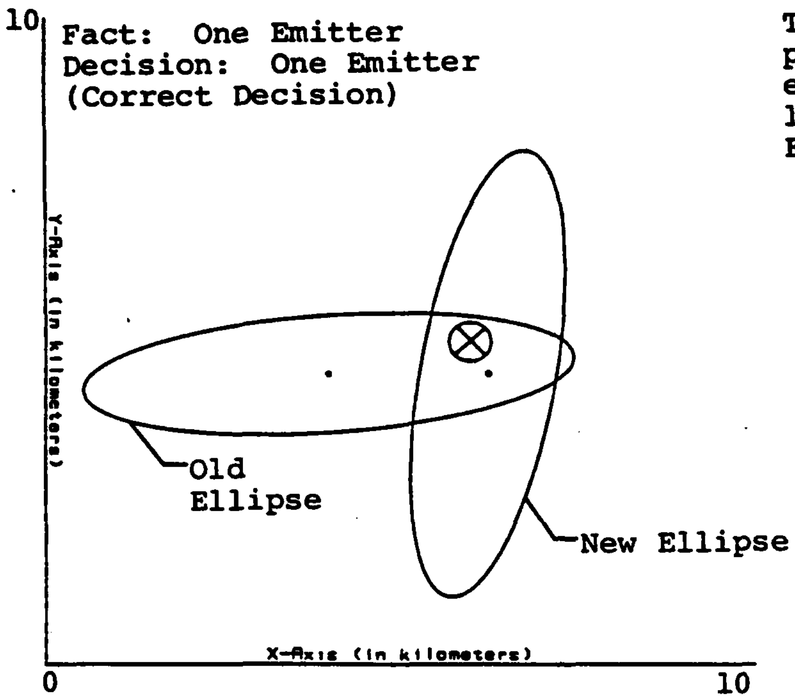


⊗ Known Emitter

○ Unknown Emitter

• Location Estimate

FIGURE 2 (Con't.)

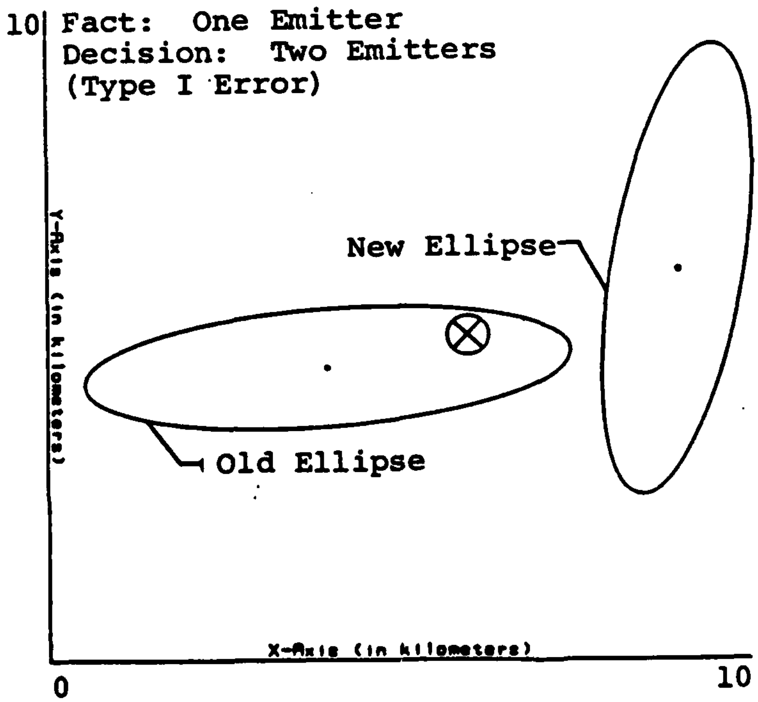


The Database contains position estimates and ellipses; the exact location of the original Emitter is not known.

Error Variances

$$\sum_{Old} = \begin{bmatrix} 2.00 & 0.15 \\ 0.15 & 0.15 \end{bmatrix}$$

$$\sum_{New} = \begin{bmatrix} 0.2 & 0.3 \\ 0.3 & 2.0 \end{bmatrix}$$



⊗ Known Emitter

No Unknown Emitter

• Location Estimate

This report addresses some of the statistical tests used in correlation of battlefield data. Based on certain assumptions (see Page 14) these tests allow one to either accept or reject that the new data correspond to a previously identified location.

SENSITIVITY OF FREQUENCIES OF TEST ERRORS TO NATURE OF INCOMING DATA

The statistical tests used in the self-correlation algorithms were designed for data in which sensor error follows the Normal, or Gaussian, probability law with known spread (variance-covariance). In addition these tests assume that the mean (expected) value of the data distribution is the true location of the emitter.

This report considers error frequency when:

1. The underlying error distributions are not Normal but follow some other type of observed distribution.

2. The variance-covariance (spread) is not known but is estimated or is incorrect.

3. The mean of the data distribution is incorrectly estimated (for example, using the sample mode in place of the sample mean when the data distribution is skewed).

The relevant questions concerning the decision errors are:

1. Can such errors occur?

2. Under what conditions can decision errors occur?

3. If such errors can occur, then at what frequency (with what probability)?

In response to the first question, errors of both type I and II can occur.

The answers to the second and third questions are the subject of this report.

II. CONCLUSION

Even when the incoming data are normally distributed with known variance-covariance, the error rates in the test decisions can cause problems when the sample sizes of the incoming data sets are small and the measurement errors are not small. When the variance-covariance of the incoming data are estimated and not known, the error rates in the test decisions become even larger.

The frequency of errors is less sensitive to changes in the underlying distribution than it is to changes in variance-covariance values and incorrect estimation of the mean of the data distribution.

Type I error (deciding the data do not belong to an indentified emitter location when in fact they do) occurs most frequently when:

1) The mean of the data distribution is incorrectly estimated (for example, using the sample mode in place of the sample mean when the data distribution is skewed).
2) The sample size is small and the variance-covariance is assumed known but is actually estimated.
3) The variance-covariance is assumed to have known values but is actually larger, possibly due to calibration error.

Type II error (deciding the data belong to an identified emitter location when they do not) occurs most frequently for:

1) small sample sizes
2) large error spread (variance-covariance)
3) underestimated error spread (assuming the variance-covariance is one set of values when it fact it is larger)

The errors discussed above are not isolated. As the flow diagrams in the Introduction indicate, these errors can cause one to update the wrong confidence ellipse, form a confidence ellipse for a phantom emitter or fail to form a confidence ellipse for an existing emitter.

The initial description of the problem, and hence the original computer work, assumed four parameters in each data observation: position coordinates plus two signal parameters. Later, however, it was discovered that there are separate tests for signal parametrics. In light of this, the computer programs were modified to use only the two position coordinates. Both sets of results are contained in the appendices. The more recent bivariate results are similar to those obtained using our original 4-variate model. Thus, conclusions drawn from the larger body of 4-variate simulation results may be carried over to the more realistic bivariate case.

In the next section of this report the error frequencies are discussed in more detail.

III. ERROR RATES OF STATISTICAL TESTS

When testing whether or not the new data are from a previously identified location certain assumptions are made concerning the new data. In particular, the self-correlation tests discussed in this report were designed for data in which the sensor error follows the Normal, or Gaussian, probability law. The goal of this casebook is to study the robustness of these self-correlation tests when these assumptions are not met (the tests are described in Appendix F).

This section will consider error rates as a function of underlying assumptions concerning the new data only. For the purpose of determining the robustness of the statistical tests under varying conditions it is sufficient to assume that the location of the "known" emitter is fixed. In other words, it is reasonable to assume that the location of the emitter in the data base is known with complete certainty. Why this is so, and how the results are applied to the actual self-correlation process, is discussed in section IV.

Throughout this report we assume that the new data are independent observations from a multivariate distribution whose mean (expected) value is the true location of an emitter. Sensitivity studies are done for the tests for several different error distributions. Further, for each of these distributions, studies are performed with varying assumptions concerning the average spread of the sensor error (assumptions concerning the variance-covariance of the data).

Test error rates are estimated for four different error distribution types, which are sketched below.

1. Normal (Gaussian). A symmetric distribution.

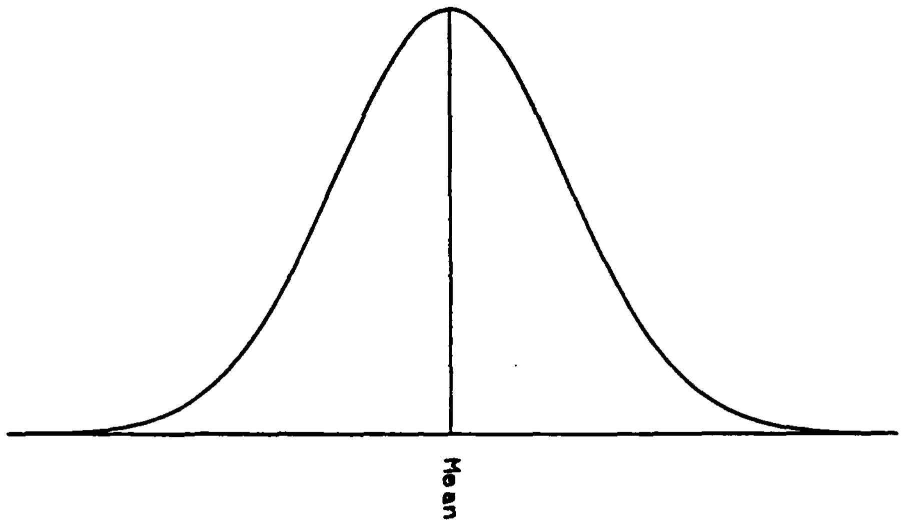
2. Gillis. A symmetric distribution with very fat tails.

3. Triangular-uniform. A symmetric distribution with tails between the Normal and the Gillis in size.

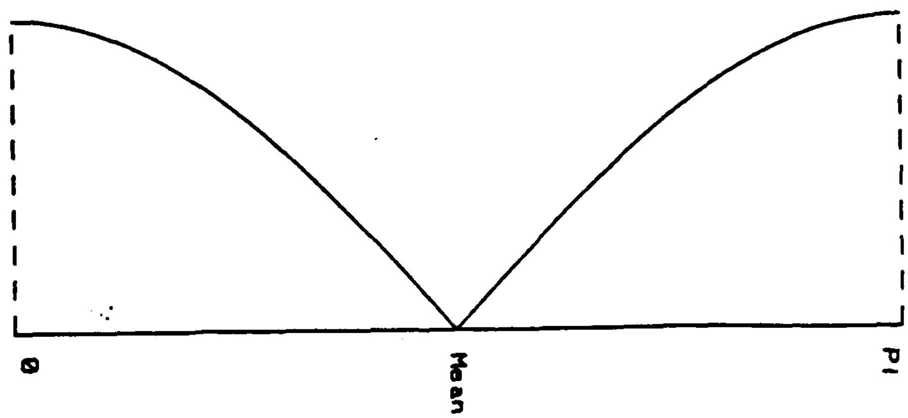
4. Gamma. A skewed (non-symmetric) distribution. Note that the mean and the mode are unequal for skewed distributions.

It has been shown that for skewed distributions the sample estimate used for emitter location can be biased away from the true location. For this casebook we attempt to simulate possible effects of the bias. For the Gamma distribution (the only skewed distribution in the casebook), two different studies are performed. The first study correctly uses the sample mean as the unbiased estimate for emitter location. The second study improperly uses the sample mode (most likely value) as the estimate for emitter location. The error involved in using the sample mode is shown graphically in Figure 3, below, for one specific bivariate case.

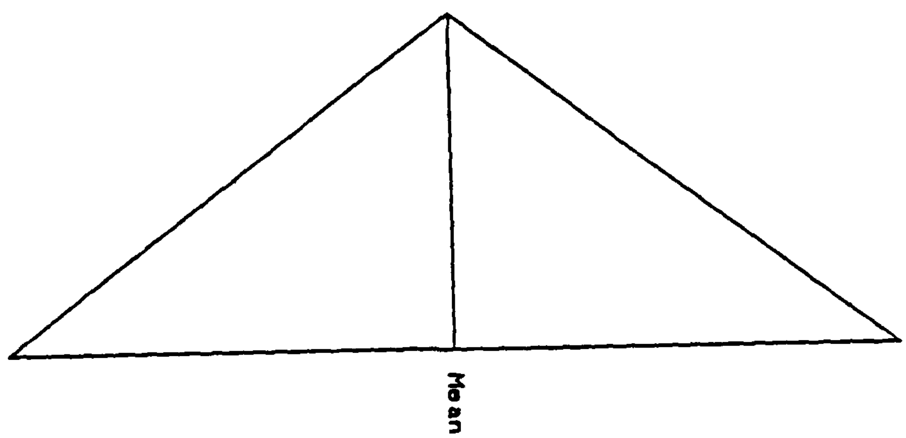
Normal Distribution



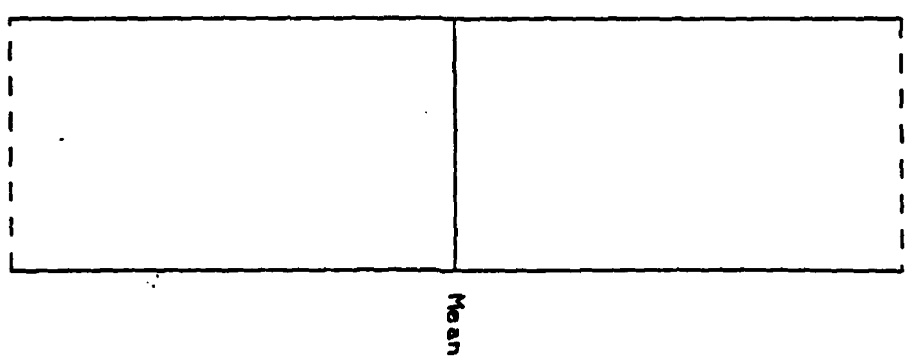
Standard Gillis Distribution



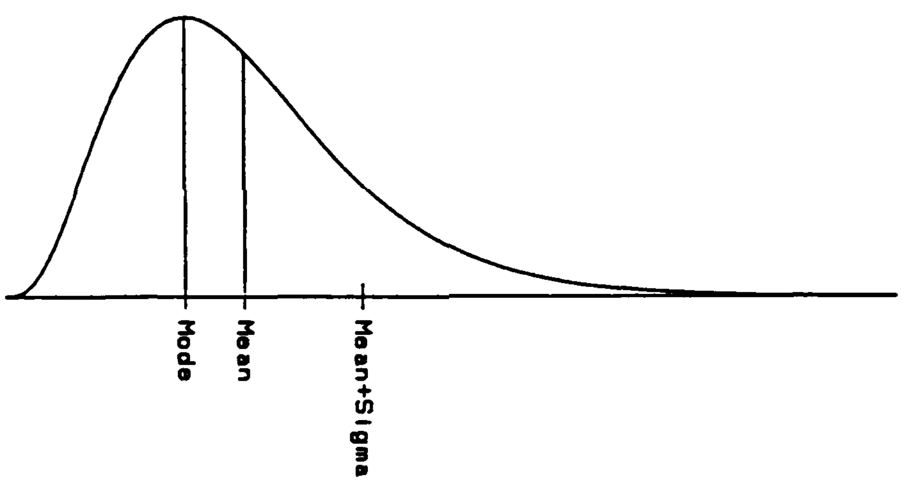
Triangular Distribution



Uniform Distribution

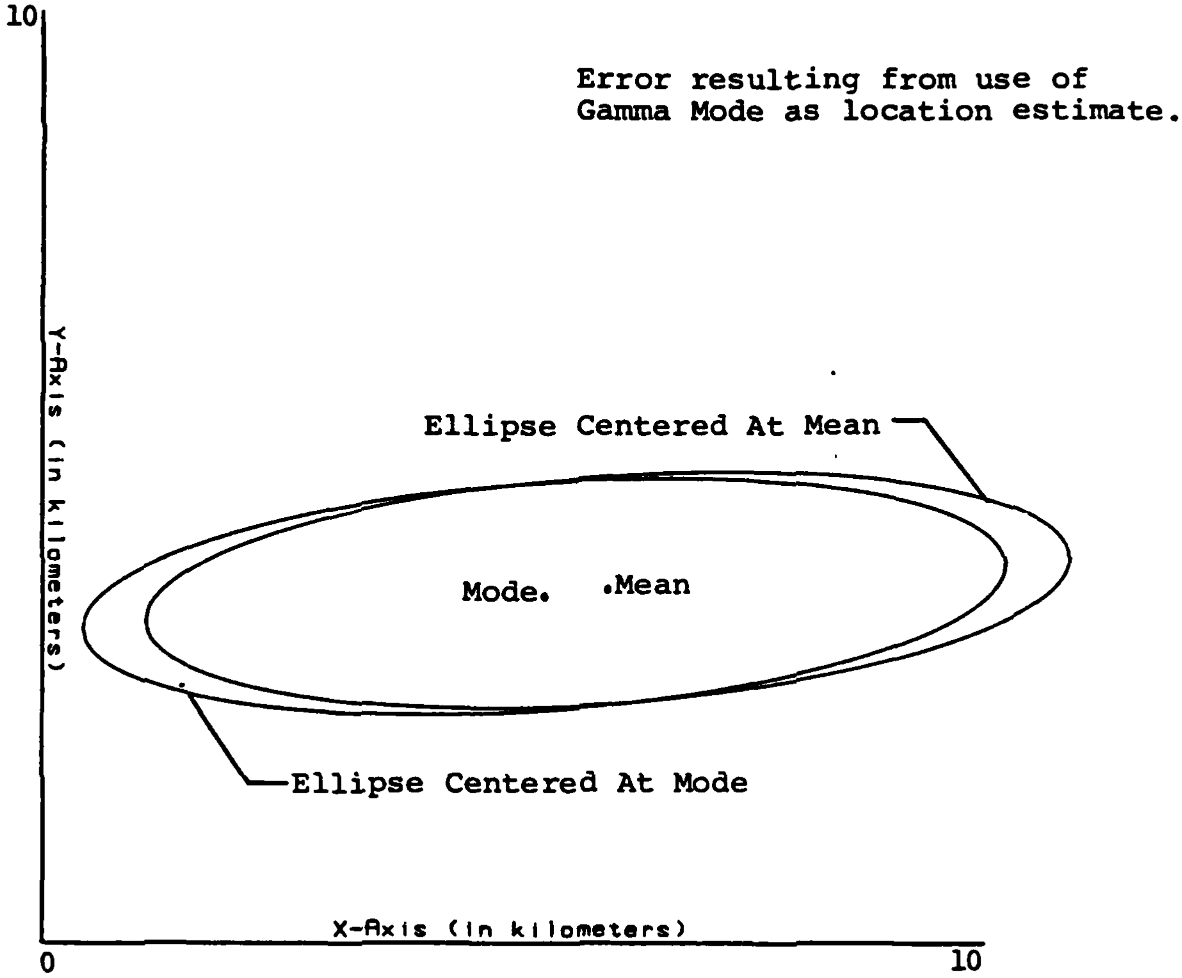


Gamma Distribution



11

FIGURE 3



$$\sum = \begin{bmatrix} 4.0 & 0.3 \\ 0.3 & 0.3 \end{bmatrix}$$

Mean: (6,4)

Mode: (5.333, 3.925)

In addition, for each of these distributions, test error rates are estimated for the following four basic assumptions about data variance-covariance (the average spread of the sensor error):

Case 1. The variance-covariance of the new data is known.

Case 2. The variance-covariance is estimated from the data.

Case 3. The variance-covariance is assumed to be known, but the estimated variance-covariance is used, i.e. no estimation error is assumed. This is an incorrect assumption when small sample sizes are used in the estimation.

Case 4. The variance-covariance is assumed known but is incorrect by a certain percentage.

The statistical tests used in each of these four cases are detailed in Appendix F.

Throughout all the cases discussed above, the sample sizes of the new data are varied in order to assess the test error rates as a function of sample size. Also the magnitude of the variance-covariance matrix is varied for each case studied.

Estimated test error rates for different distributions and assumptions about the variance-covariance of the data were obtained by simulation. From 200 to 500 data sets were simulated for each distribution, assumed variance-covariance, sample size and location. Simulation methods are detailed in Appendix G.

The following table summarizes the results of the simulations. This summary compares the errors obtained by varying the assumptions (distributions, variance-covariance) to those errors obtained when all assumptions underlying the tests are met. Even if all assumptions underlying the statistical tests are not met the tests may be robust with respect to these unmet assumptions, i.e. they may yield errors comparable to those obtained when all the assumptions are met.

Table 1 contains classification as to Robust or Not Robust for those distributions and variance-covariance cases simulated. If Type I and/or Type II errors are greater than those expected under the correct assumptions then the cases are classified as Not Robust. For further details see the simulation results listed under each distribution type.

TABLE 1

Case	Normal	Gillis	Triangular-uniform	Gamma*	Gamma Mode**
(1) Variance-covariance known	Assumptions met	Robust	Not Robust	Not Robust	Not Robust
(2) Variance-covariance estimated	Assumptions met	Not Robust	Not Robust	Not Robust	Not Robust
(3) Variance-covariance assumed known but estimated	Not Robust	Not Robust	Not Robust	Not Robust	Not Robust
(4) Variance-covariance assumed known but in error by a percentage.	Not Robust	Not Robust	Not Robust	Not Robust	Not Robust

* These Gamma distributions are Gamma distributions where the sample mean is used to estimate the emitter location or parametrics.

** These Gamma distributions are Gamma distributions where the sample mode (most frequent data values) are used to estimate the emitter location or parametrics.

All studies were performed by setting the test probability of type I error (probability of deciding the new data are not from the identified location when in fact they are) at 0.05.

The following five tables contain summary results by distribution. For each variance-covariance case, probabilities of type I and type II errors are discussed. More complete results are given in the appendicies.

TABLE 2: NORMAL DISTRIBUTION

CASE	PROBABILITY OF TYPE I ERROR	PROBABILITY OF TYPE II ERROR
(1) Variance-covariance known	Set at 0.05	Varies between 0.90 and 0 depending on sample size, assumed difference in location of emitters and magnitude of variance-covariance
(2) Variance-covariance estimated from data	Set At 0.05	Varies between 0.94 and 0 depending on sample size, assumed difference in location of emitters and magnitude of variance-covariance
(3) Variance-covariance assumed known but estimated	Varies between 0.77 and 0.05 depending on sample size. Larger error rates occur for 4-variate data	Varies between 0.52 and 0. Larger error rates occur for bivariate data
(4) Variance-covariance assumed known but in error by a percentage	Varies between .22 and 0 depending on the magnitude of variance-covariance error	Varies between 0.97 and 0 depending on sample size, assumed difference in location of emitters and magnitude of variance-covariance error

In all cases as either the sample size increases or the distance between the emitters increases, the probability of Type II error goes to 0.

TABLE 3: GILLIS DISTRIBUTION

CASE	PROBABILITY OF TYPE I ERROR	PROBABILITY OF TYPE II ERROR
(1) Variance-covariance known	Agrees with the set level of 0.05	Varies between 0.71 and 0 depending on sample size, assumed difference in location of emitters and magnitude of variance-covariance
(2) Variance-covariance estimated from data	Varies between 0.12 and 0.03	Varies between 0.94 and 0 depending on sample size, assumed difference in location of emitters and magnitude of variance-covariance
(3) Variance-covariance assumed known but estimated	Varies between 0.70 and 0.04 depending on sample size. Error rates are higher for 4-variate data	Varies between 0.68 and 0. Error rates are higher for bivariate data
(4) Variance-covariance assumed known but in error by a percentage	Varies between 0.24 and 0 depending on the magnitude of variance-covariance error	Varies between 0.92 and 0 depending on sample size, assumed difference in location of emitters and magnitude of variance-covariance error

In all cases as either the sample size increases or the distance between the emitters increases, the probability of Type II error goes to 0.

TABLE 4: TRIANGULAR-UNIFORM DISTRIBUTION

CASE	PROBABILITY OF TYPE I ERROR	PROBABILITY OF TYPE II ERROR
(1) Variance-covariance known	Agrees with the set level of 0.05	Varies between 0.85 and 0 depending on sample size, assumed difference in location of emitters and magnitude of variance-covariance*
(2) Variance-covariance estimated from data	Agrees with the set level of 0.05	Varies between 0.95 and 0 depending on sample size, assumed difference in location of emitters and magnitude of variance-covariance*
(3) Variance-covariance assumed known but estimated	Varies between 0.75 and 0.05 depending on sample size. Higher error rates occur for 4-variate data	Varies between 0.65 and 0. Higher error rates occur for bivariate data
(4) Variance-covariance assumed known but in error by a percentage	Varies between 0.24 and 0 depending on the magnitude of variance-covariance error	Varies between 0.94 and 0 depending on sample size, assumed difference in location of emitters and magnitude of variance-covariance error

In all cases as either the sample size increases or the distance between the emitters increases, the probability of Type II error goes to 0.

* Error is greater than when the distribution is Normal

TABLE 5: GAMMA DISTRIBUTION

CASE	PROBABILITY OF TYPE I ERROR	PROBABILITY OF TYPE II ERROR
(1) Variance-covariance known	Agrees with the set level of 0.05	Varies between 0.76 and 0 depending on sample size, assumed difference in location of emitters and magnitude of variance-covariance*
(2) Variance-covariance estimated from data	Varies from 0.12 to 0.03. Small sample sizes for bivariate data have the highest error rate	Varies between 0.92 and 0 depending mainly on sample size, assumed difference in location of emitters and magnitude of variance-covariance*
(3) Variance-covariance assumed known but estimated	Varies between 0.80 and 0.06 depending on sample size. Higher error rates occur for 4-variate data	Varies between 0.42 and 0 depending mainly on sample size, assumed difference in location of emitters and magnitude of variance-covariance
(4) Variance-covariance assumed known but in error by a percentage	Varies between 0.24 and 0.02 depending on the magnitude of variance-covariance error	Varies between 0.96 and 0 depending on sample size, assumed difference in location of emitters and magnitude of variance-covariance error

In all cases as either the sample size increases or the distance between the emitters increases, the probability of Type II error goes to 0.

* Error is greater than when the distribution is Normal

TABLE 6: GAMMA MODE DISTRIBUTION

CASE	PROBABILITY OF TYPE I ERROR	PROBABILITY OF TYPE II ERROR
(1) Variance-covariance known	Varies between 1.0 and 0.88 depending on sample size. The larger the sample size the larger the probability of error	Almost 0
(2) Variance-covariance estimated from data	Varies between 1.0 and 0.09 depending on sample size. The larger the sample size the larger the probability of error	Varies between 0.88 for small sample size and 0 for larger sample size. Also depends on assumed difference in location of emitters*
(3) Variance-covariance assumed known but estimated	Varies between 1.0 and 0.93 depending mainly on sample size. The larger the sample size the larger the probability of error	Almost 0
(4) Variance-covariance assumed known but in error by a percentage	Varies between 1.0 and 0.76 depending on sample size and magnitude of the variance-covariance error	From 0.24 to 0.

In all cases the probability of Type I error increases as the sample size increases.

* Error is greater than when the distribution is Normal

IV. CONFIDENCE ELLIPSE CONSIDERATIONS

In practice, "known" emitter locations are stored in a database as a location estimate and a confidence ellipse. That is, while the emitter is "known" to the analysis center, its actual position is uncertain. The self-correlation statistical tests take this into account. However, for the purpose of this analysis it is reasonable to assume that the location of the "known" emitter is in fact known with certainty. The results obtained under this assumption may be applied to the actual situation with complete generality. This section will provide a basis for this statement, and explain how to apply the results to the more complex (two ellipse) case, which is illustrated in Figure 2 in the Introduction. Additional ellipse considerations are the effect of sample size and the effect of inaccurate covariance matrices.

1. Two Ellipse Case

The statistical test uses two location estimates and a covariance matrix for each. However, a comprehensive study involving many pairs of covariance matrices would have too many cases to report. Consequently, it is important to find classes of cases that yield the same result. Here is an example of one such class. The following three pairs of covariance matrices yield the same value of the test statistic for any given pair of location estimates (Note that the following arguments hold for 4-variate covariance matrices as well).

	Covariance for Estimate 1	Covariance for Estimate 2	Sum of Matrices
Pair 1)	$\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}$	$\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$	$\begin{bmatrix} 5 & 2 \\ 2 & 6 \end{bmatrix}$
Pair 2)	$\begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$	$\begin{bmatrix} 5 & 2 \\ 2 & 6 \end{bmatrix}$
Pair 3)	$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 5 & 2 \\ 2 & 6 \end{bmatrix}$	$\begin{bmatrix} 5 & 2 \\ 2 & 6 \end{bmatrix}$

These three pairs yield the same value of the test statistic because the formula is based on <u>the sum of the covariance matrices</u>. Since each pair of matrices above sum to the same matrix, each is equivalent where the

31

statistical test is concerned. The third pair is different from the first two, however, in that there is no uncertainty in estimate 1. This is the same as knowing the precise location of one emitter. Note that this special case may easily be found for any general two ellipse case by simply taking the sum of the covariances of the two ellipses and using the sum in the statistical test.

The statistical test cannot differentiate between any of the three pairs of covariance matrices listed above. Consequently, it is clear that by examining the special case above one produces results which apply to the first two pairs of matrices. Indeed, the results will apply to any pair of covariance matrices which sum to the second matrix in Pair 3. Further, the results will not vary between the cases because of skewness, kurtosis, bias and so on, though uncertainty in the covariance matrix may have an effect.

Appendices A through E contain the results of the simulations in tabular form. At the top of each table is the "special case" matrix which was used to simulate the data. By the reasoning above, this table also applies to any two ellipse case in which the covariance matrices sum to the matrix shown on the table. Note that this "special case" matrix is also referred in this report as the composite matrix.

The test used in Cases 2 and 3 is slightly different. In it, the formula for the test statistic uses the sum of two estimated covariance matrices, and the correct test statistic is based on the F distribution. While the results for Cases 2 and 3 may also be generalized in the manner shown above, the process itself is somewhat different. It is also slightly more complex, and since we have not encountered the F-test in practice we will not explain it here. Suffice it to say that the results listed in the appendices for Cases 2 and 3 are an accurate assessment of the robustness of the F-test.

2. Effect of Sample Size on Covariance Matrices

In the simulations run for this casebook, the data were generated as observations from a number of different multivariate distributions, such as the Normal and the Gamma. The statistical tests assume that as the sample size increases our uncertainty as to the emitter location (size of the covariance matrix) decreases. In the simple or "special" case, typified by "Pair 3" in the example in part 1 of this section, if the covariance matrix is the known (or estimated) error owing to sensor inaccuracy then the statistical test uses this covariance matrix divided by the appropriate sample size, n. For example, if the sensor covariance matrix is

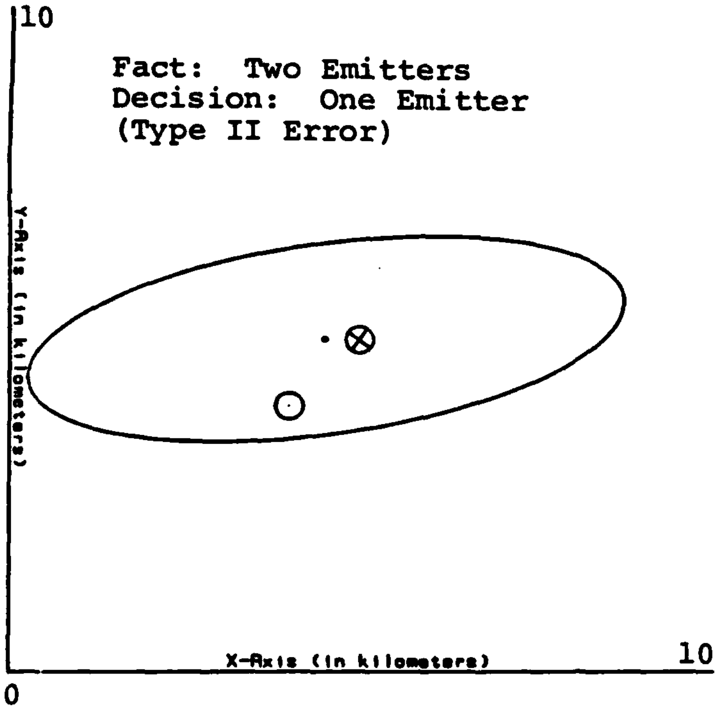
$$\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$$

and sample size n = 2, then the matrix used in summing the composite matrix is

$$\begin{bmatrix} 1 & 1/2 \\ 1/2 & 2 \end{bmatrix}$$

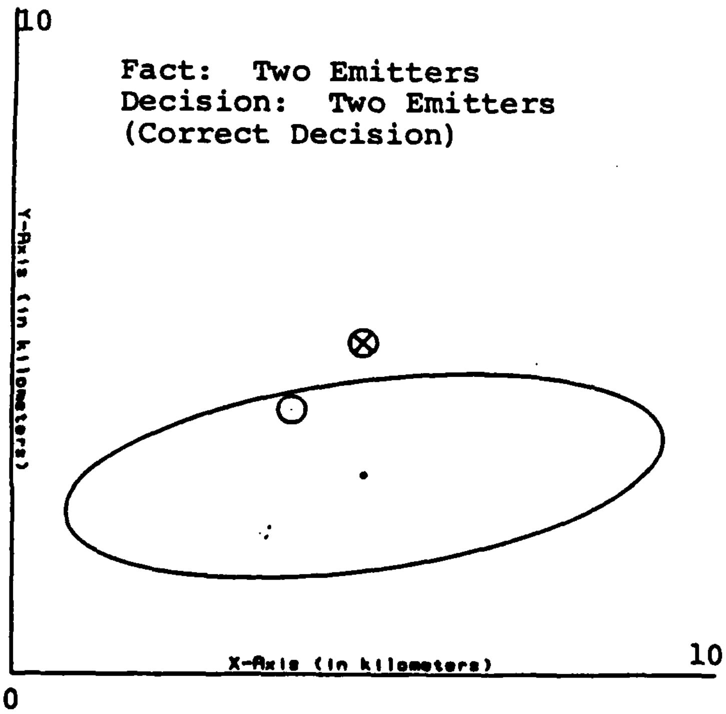
FIGURE 4

CASE 1



Error Variance

$$\sum = \begin{bmatrix} 4.0 & 0.3 \\ 0.3 & 0.3 \end{bmatrix}$$



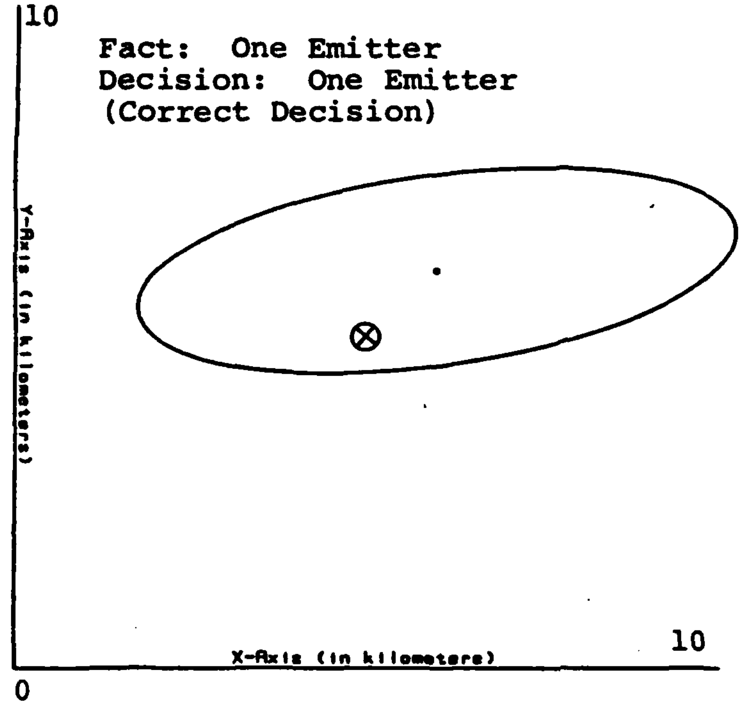
⊗ Known Emitter

○ Unknown Emitter

• Location Estimate

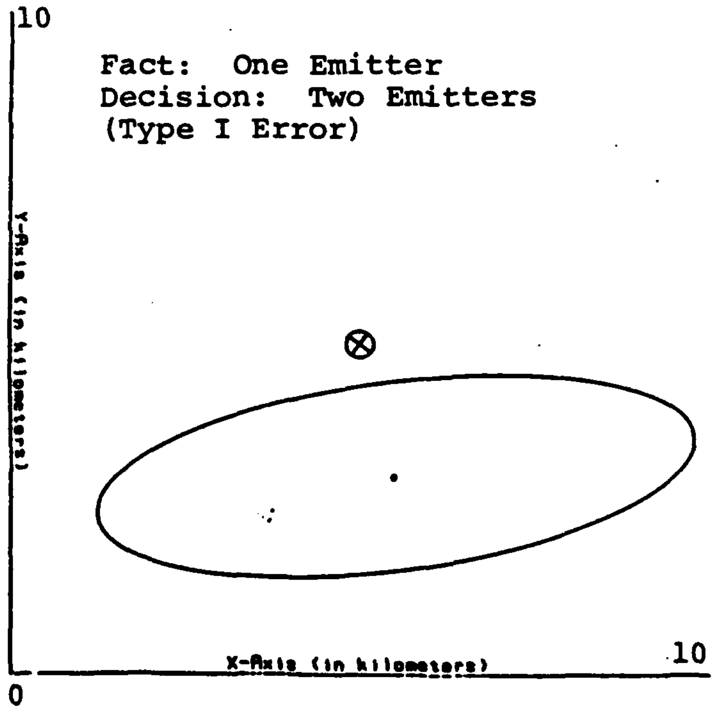
FIGURE 4 (Con't.)

CASE 1



Error Variance

$$\Sigma = \begin{bmatrix} 4.0 & 0.3 \\ 0.3 & 0.3 \end{bmatrix}$$



⊗ Known Emitter

No Unknown Emitter

• Location Estimate

In order to study the effect of sample size on the accuracy of the statistical tests we bring sample size into the calculations by simply dividing it into the composite matrix. In practice there would be two sample sizes; one for the estimate in the database and one for the new data. Time does not permit considering all possible cases. Therefore the effect of sample size is studied by simply considering the frequency of test errors where the composite matrix given at the top of each table (in the Appendices) is divided by the sample size noted in the table. In all of the cases shown, the test matrix is quite small when the sample size increases to 100, i.e. 1/100 of the original covariance matrix.

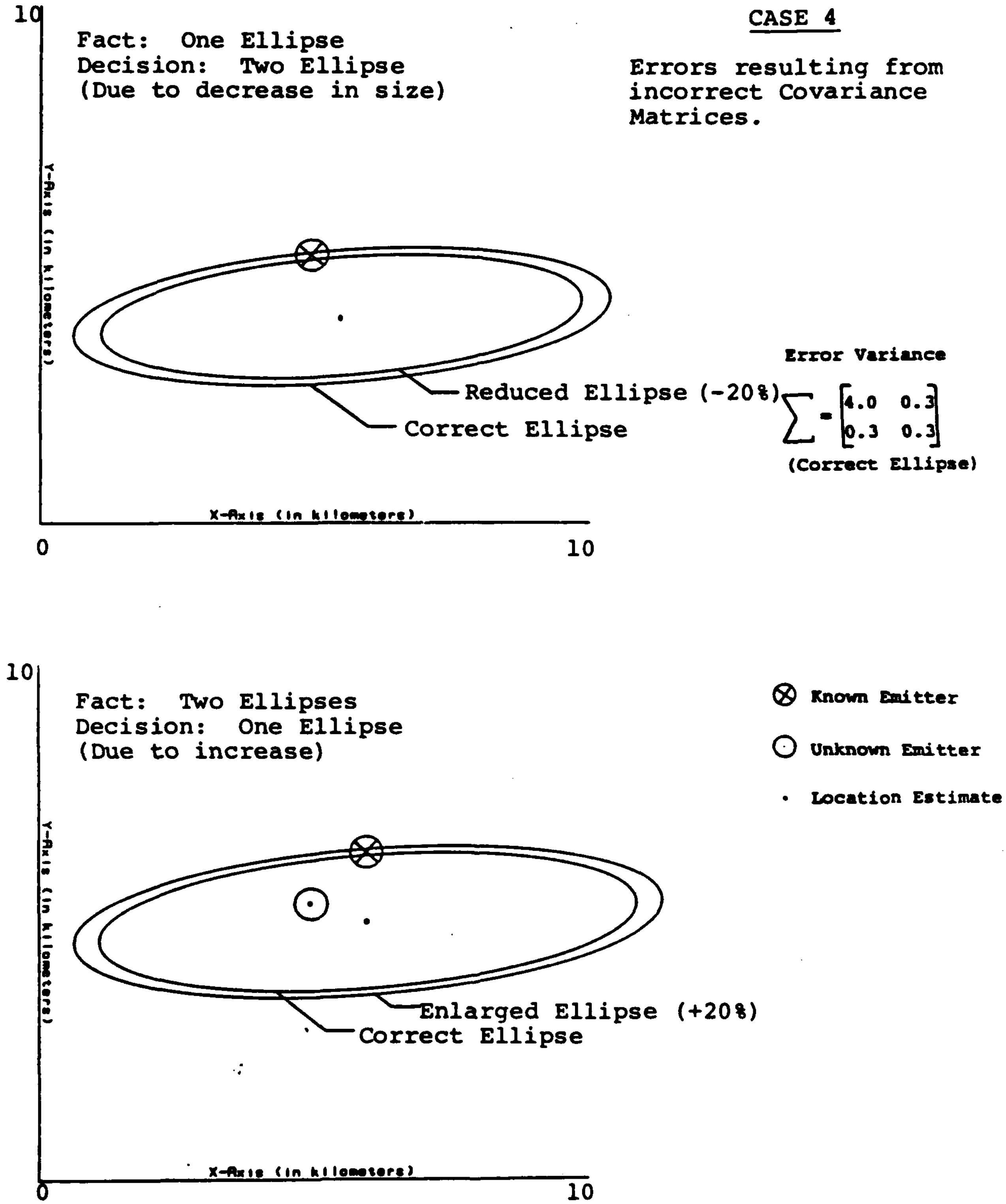
It should be noted that the above discussion of sample size depends on the classic statistical methods of estimation. Sample size as used here is not directly applicable to the methods used in the actual fix algorithms we have studied. That is, it is not a well-defined concept in the context of position-fixing algorithms. However, this does not mean that it is not a useful concept in this setting. The sample size used in the Appendices may be thought of as measuring the amount or goodness of the information that was actually used in calculating the position estimate.

3. The Effect of Inaccurate Covariance Matrices

Though the following issue is not directly related to the concerns in the preceding paragraphs, it will be discussed here because it involves the composite covariance matrix used in the chi-square statistical test. This issue is the error that can occur when the data covariance matrices used in calculating the composite matrix are not the true covariance matrices. In this analysis (represented by Case 4 in the appendices) the simplifying assumption of one emitter location known will be used. Further, the incorrect covariance matrix used will only be off by a small percentage of the correct covariance matrix. The purpose of this is to show the effect of small errors. The covariance matrix was varied by amounts ranging from -4% to 4%.

Figure 5 shows two of the possible errors resulting from the incorrect covariance matrices. In both of the cases shown, an incorrect decision was made because the confidence ellipse was too large or too small. If the correct matrix had been used, the correct decision would have been made. Note that the matrices used in Figure 5 were increased or decreased by 20%. Such a large factor was used so that a clear graph could be produced. In fact, an increase or decrease of as little as 1% or 2% is enough to cause a significant number of incorrect decisions.

FIGURE 5

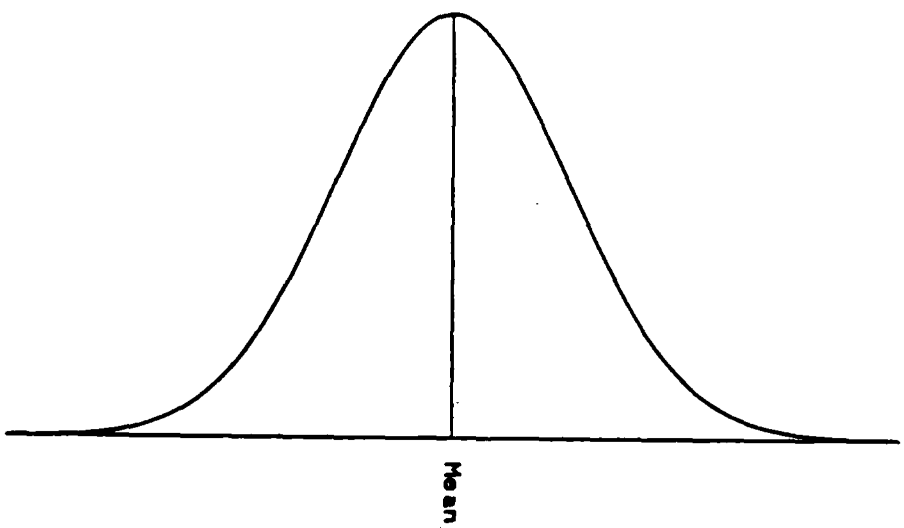

CASE 4
Errors resulting from incorrect Covariance Matrices.
Fact: One Ellipse
Decision: Two Ellipse
(Due to decrease in size)
Reduced Ellipse (-20%)
Correct Ellipse
Error Variance
$\Sigma = \begin{bmatrix} 4.0 & 0.3 \\ 0.3 & 0.3 \end{bmatrix}$
(Correct Ellipse)
Y-Axis (in kilometers)
X-Axis (in kilometers)
0
10
10
Fact: Two Ellipses
Decision: One Ellipse
(Due to increase)
Enlarged Ellipse (+20%)
Correct Ellipse
Known Emitter
Unknown Emitter
Location Estimate
Y-Axis (in kilometers)
X-Axis (in kilometers)
0
10
10

APPENDIX A

Normal Distribution

The Normal Distribution is the underlying distribution assumed for the data in the tests studied. The Normal distribution is often assumed to be the distribution of equipment measuring error.

Sample sizes vary from 5 to 50. The tables which follow are categorized by the four basic assumptions about the data variance-covariance (see Page 14). Both 4-variate and bivariate cases are listed.



Normal Distribution, Case 1

The test assumes the underlying distribution is Normal, that the variance-covariance matrix for the distribution is known, and that the location is estimated by the mean of the data.

All assumptions are met.

Summary

Probability of Type I error:

Set at 0.05

Probability of Type II error:

Varies between 0.90 and 0 depending on sample size, assumed difference in location of emitters and magnitude of variance-covariance.

TABLE A-1

$$\sum = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

CASE 1

NORMAL

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

Sample Size

DIFFERENCE IN LOCATION OF EMITTERS	5	10	20	50	100
0.0	96	93	95	95	-
0.5	60	29	5	0	-
1.0	5	0	0	0	-
1.5	0	0	0	0	-
2.0	0	0	0	0	-
2.5	0	0	0	0	-
3.0	0	0	0	0	-

When the difference in location is 0 the estimated probability of Type I error is

$$1- \frac{\text{Percentage}}{100}$$

Thus for Sample Size 10 the estimated probability of Type I error is

$$1 - \frac{93}{100} = 0.07.$$

The estimated probability of Type II is simply the

$$\frac{\text{Percentage}}{100}$$

Thus for Sample Size 10 with a difference in location of emitters of 0.5 the probability of Type II error is

$$\frac{29}{100} = 0.29.$$

When the difference in location is given as .5 this means that the difference is .5 in each of the "directions".

TABLE A-2

$$\sum = \begin{bmatrix} 3 & 1 & 1 & .5 \\ 1 & 2 & .3 & .2 \\ 1 & .3 & 4 & 1 \\ .5 & .2 & 1 & 2 \end{bmatrix}$$

CASE 1

NORMAL

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS

	Sample Size				
	5	10	20	50	100
0.0	95	95	95	94	-
0.5	90	77	61	23	-
1.0	62	28	5	0	-
1.5	24	1	0	0	-
2.0	4	0	0	0	-
2.5	1	0	0	0	-
3.0	0	0	0	0	-

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1- \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE A-3

$$\sum = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

CASE 1

NORMAL

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size				
	5	10	20	50	100
(.5,.5,.5,.5)	60	28	4	0	0
(.5,.5,.5,1)	35	6	0.1	0	0
(.5,.5,1,1)	20	1	0	0	0
(.5,1,1,1)	8	0.1	0	0	0
(1,1,1,1)	4	0	0	0	0
(1,1,1,1.5)	1	0	0	0	0

$$\text{Estimated probability of Type II error} = \frac{\text{Percentage}}{100}$$

TABLE A-4

$$\sum = \begin{bmatrix} 3 & 1 & 1 & .5 \\ 1 & 2 & .3 & .2 \\ 1 & .3 & 4 & 1 \\ .5 & .2 & 1 & 2 \end{bmatrix}$$

CASE 1

NORMAL

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
(.5,.5,.5,.5)	88	80	61	19	1.2
(1,1,1,1)	61	29	4	0	0
(1,5,1.5,1.5,1.5)	24	2	0	0	0
(2,2,2,2)	4	.02	0	0	0
(2.5,2.5,2.5,2.5)	.2	0	0	0	0
(3,3,3,3)	0	0	0	0	0

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$

TABLE A-5

$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

CASE 1 - Bivariate

NORMAL

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	94	96	95	94	96
0.5	76	46	14	0	0
1.0	18	2	0	0	0
1.5	1	0	0	0	0
2.0	0	0	0	0	0
2.5	0	0	0	0	0
3.0	0	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I Error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II Error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

Normal Distribution, Case 2

The test assumes the underlying distribution is Normal, that the variance-covariance matrix is estimated from the data, and that the location is estimated by the mean of the data.

All assumptions are met.

Summary

Probability of Type I error:

Set at 0.05

Probability of Type II error:

Varies between 0.94 and 0 depending on sample size, assumed difference in location of emitters and magnitude of variance-covariance.

TABLE A- 6

$$\sum = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

CASE 2

NORMAL

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size				
	5	10	20	50	100
0.0	96	94	94	95	-
0.5	90	65	9	0	-
1.0	90	7	0	0	-
1.5	84	0	0	0	-
2.0	76	0	0	0	-
2.5	74	0	0	0	-
3.0	66	0	0	0	-

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1- \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE A- 7

$$\sum = \begin{bmatrix} 3 & 1 & 1 & .5 \\ 1 & 2 & .3 & .2 \\ 1 & .3 & 4 & 1 \\ .5 & .2 & 1 & 2 \end{bmatrix}$$

CASE 2

NORMAL

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size				
	5	10	20	50	100
0.0	96	95	96	97	-
0.5	94	88	67	21	-
1.0	91	62	8	0	-
1.5	87	30	2	0	-
2.0	90	6	0	0	-
2.5	84	0	0	0	-
3.0	82	0	0	0	-

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE A- 8

$$\sum = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

CASE 2-Bivariate

NORMAL

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

Sample Size

DIFFERENCE IN LOCATION OF EMITTERS	5	10	20	50	100
0.0	92	95	98	97	96
0.5	84	67	26	0	0
1.0	62	14	0	0	0
1.5	35	0	0	0	0
2.0	12	0	0	0	0
2.5	2	0	0	0	0
3.0	1	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Note that two sets of simulations were run for Sample Size 5.

Estimated probability of Type I Error = $1-\frac{\text{Percentage}}{100}$

when difference in location is 0.

Estimated probability of Type II Error = $\frac{\text{Percentage}}{100}$

when difference in location is not 0.

TABLE A-9

$$\sum = \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix}$$

CASE 2-Bivariate

NORMAL

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	95	99	96	96	92
0.5	86	56	24	0	0
1.0	60	8	0	0	0
1.5	30	0	0	0	0
2.0	13	0	0	0	0
2.5	4	0	0	0	0
3.0	1	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Note that two sets of simulations were run for Sample Size 5.

Estimated probability of Type I Error = $1 - \frac{\text{Percentage}}{100}$

when difference in location is 0.

Estimated probability of Type II Error = $\frac{\text{Percentage}}{100}$

when difference in location is not 0.

Normal Distribution, Case 3

The test assumes the underlying distribution is Normal, that the variance-covariance matrix is known, and the location is estimated by the mean of the data.

The first and third assumptions are met. However, the variance-covariance matrix used in the test is estimated from the data.

Summary

Probability of Type I error:

Varies between 0.77 and 0.05 depending on sample size. Larger error rates occur for 4-variate data.

Probability of Type II error:

Varies between 0.52 and 0. Larger error rates occur for bivariate data.

TABLE A-10

$$\sum = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

CASE 3

NORMAL

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

Sample Size

DIFFERENCE IN LOCATION OF EMITTERS	5	10	20	50	100
0.0	28	72	86	95	-
0.5	8	12	4	0	-
1.0	0.5	0	0	0	-
1.5	0	0	0	0	-
2.0	0	0	0	0	-
2.5	0	0	0	0	-
3.0	0	0	0	0	-

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error $= 1- \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error $= \frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE A-11

$$\sum = \begin{bmatrix} 3 & 1 & 1 & .5 \\ 1 & 2 & .3 & .2 \\ 1 & .3 & 4 & 1 \\ .5 & .2 & 1 & 2 \end{bmatrix}$$

CASE 3

NORMAL

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	24	76	91	94	-
0.5	18	59	45	17	-
1.0	10	10	4	0	-
1.5	3	1	0	0	-
2.0	2	0	0	0	-
2.5	0	0	0	0	-
3.0	0	0	0	0	-

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE A - 12

$$\sum = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

CASE 3-Bivariate

NORMAL

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	75	87	89	92	91
0.5	48	38	20	2	0
1.0	14	3	0	0	0
1.5	2	0	0	0	0
2.0	0	0	0	0	0
2.5	0	0	0	0	0
3.0	0	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Note that two sets of simulations were run for Sample Size 5.

Estimated probability of Type I Error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II Error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE A-13

$$\Sigma = \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix}$$

CASE 3-Bivariate

NORMAL

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	78	86	92	94	95
0.5	52	44	23	0	0
1.0	14	1	0	0	0
1.5	2	0	0	0	0
2.0	0	0	0	0	0
2.5	0	0	0	0	0
3.0	0	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Note that two sets of simulations were run for Sample Size 5.

Estimated probability of Type I Error = $1 - \frac{\text{Percentage}}{100}$

when difference in location is 0.

Estimated probability of Type II Error = $\frac{\text{Percentage}}{100}$

when difference in location is not 0.

Normal Distribution, Case 4

The test assumes the underlying distribution is Normal, that the variance-covariance matrix is known, and that the location is estimated by the mean of the data.

The first and third assumptions are met. However, the variance-covariance matrix used in the test is varied from the true by a percentage (-4% to 4%). This could be thought of as an effect of calibration error.

Summary

Probability of Type I error:

Varies between 0.22 and 0 depending on the magnitude of the variance-covariance error.

Probability of Type II error:

Varies between 0.97 and 0 depending on sample size, assumed difference in location of emitters and magnitude of the variance-covariance error.

TABLE A-14

CASE 4

$$\sum = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

NORMAL

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

% Change in Covaraince Matrix	Diff. in Loc. Of Emitter	Sample Size 5	10	20	50	100
-4%	0	83	80	79	77	-
	0.5	38	10	0.5	0	-
	1.0	0.5	0	0	0	-
	2.0	0	0	0	0	-
-3%	0	86	85	83	85	-
	0.5	40	10	2	0	-
	1.0	2	0	0	0	-
	2.0	0	0	0	0	-
-2%	0	86	89	84	92	-
	0.5	38	21	2	0	-
	1.0	2	0	0	0	-
	2.0	0	0	0	0	-
2%	0	98	100	98	98	-
	0.5	74	43	3	0	-
	1.0	10	0	0	0	-
	2.0	0	0	0	0	-
3%	0	97	99	98	98	-
	0.5	78	46	10	0	-
	1.0	10	0	0	0	-
	2.0	0	0	0	0	-
4%	0	99	100	99	100	-
	0.5	78	44	8	0	-
	1.0	10	0	0	0	-
	2.0	0	0	0	0	-

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$

For 0% change see Table A-1 results.

TABLE A-15

CASE 4

$$\sum = \begin{bmatrix} 3 & 1 & 1 & .5 \\ 1 & 2 & .3 & .2 \\ 1 & .3 & 4 & 1 \\ 5 & 2 & 1 & 2 \end{bmatrix}$$

NORMAL

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

% Change in Covaraince Matrix

% Change	Diff. in Loc. Of. Emitter	Sample Size 5	10	20	50	100
-4%	0	84	81	78	78	-
	0.5	62	52	30	4	-
	1.0	34	8	0.5	0	-
	2.0	1	0	0	0	-
-3%	0	86	83	84	82	-
	0.5	72	56	39	7	-
	1.0	36	14	0	0	-
	2.0	1	0	0	0	-
-2%	0	88	88	91	92	-
	0.5	78	69	40	15	-
	1.0	38	19	4	0	-
	2.0	4	0	0	0	-
2%	0	99	98	97	97	-
	0.5	97	88	73	30	-
	1.0	73	36	8	0	-
	2.0	0.9	0.5	0	0	-
3%	0	99	100	98	98	-
	0.5	97	92	82	30	-
	1.0	74	40	8	0	-
	2.0	8	0	0	0	-
4%	0	98	99	100	99	-
	0.5	96	92	80	39	-
	1.0	84	50	11	0	-
	2.0	14	0	0	0	-

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$

For 0% change see Table A-2 results.

TABLE A-16

CASE 4-Bivariate

$$\sum = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

NORMAL

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

% Change in Covariance Matrix: % Change	Diff. in Loc. Of. Emitter	Sample Size: 5	10	20	50	100
-4%	0	78	86	85	87	86
	0.5	44	26	4	0	0
	1.0	5	0	0	0	0
	2.0	0	0	0	0	0
-3%	0	89	85	88	84	87
	0.5	55	30	8	0	0
	1.0	12	0	0	0	0
	2.0	0	0	0	0	0
-2%	0	91	93	93	94	92
	0.5	54	38	10	1	0
	1.0	12	2	0	0	0
	2.0	0	0	0	0	0
2%	0	98	98	96	98	96
	0.5	76	60	30	0	0
	1.0	32	1	0	0	0
	2.0	0	0	0	0	0
3%	0	96	99	96	97	98
	0.5	82	66	32	0	0
	1.0	32	3	0	0	0
	2.0	0	0	0	0	0
4%	0	99	98	96	98	100
	0.5	82	66	36	2	0
	1.0	36	4	0	0	0
	2.0	0	0	0	0	0

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$

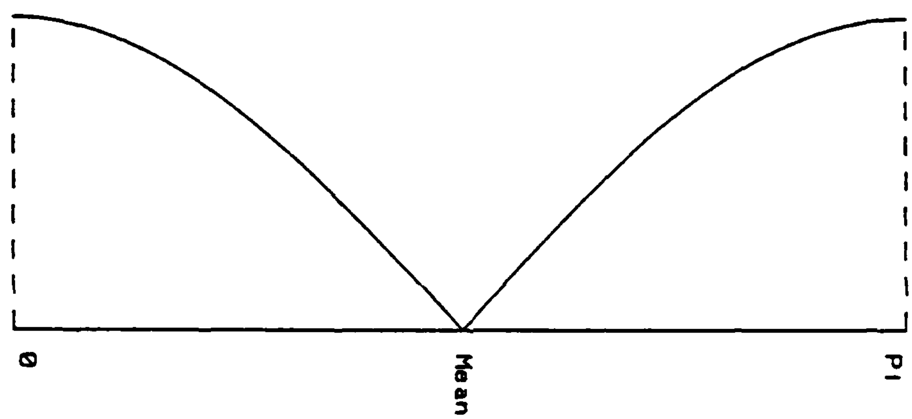
Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$

APPENDIX B

Gillis Distribution

In the simulation results which follow, the test distribution is the Gillis distribution.

The Gillis distribution is a very fat-tailed distribution based on the cosine function.



The standard univariate Gillis sketched above has the following density function.

$$f(x) = \begin{cases} \cos(x)/2 & 0 \le x \le \pi/2 \\ \cos(\pi - x)/2 & \pi/2 \le x \le \pi \end{cases}$$

For the 4-variate cases, the data are vectors with four components, each of which is an independent Gillis random variable. For the bivariate case, the data are ordered pairs, the two components of which may be dependent.

Sample sizes vary from 5 to 100. The tables which follow are categorized by the four basic assumptions about the data variance-covariance (see Page 14).

Gillis Distribution, Case 1

The test assumes the underlying distribution is Normal, that the variance-covariance matrix for the distribution is known, and that the location is estimated by the mean of the data.

The first assumption is not met. The data follow a Gillis distribution, not the Normal distribution. However, the second and third assumptions are met.

Summary

Probability of Type I error:

Agrees with the set level of 0.05.

Probability of Type II error:

Varies between 0.71 and 0 depending on sample size, assumed difference in location of emitters and magnitude of variance-covariance.

TABLE B-1

$$\sum = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

CASE 1

GILLIS

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size				
	5	10	20	50	100
0.0	96	96	95	94	95
0.5	58	30	3	0	0
1.0	3	0	0	0	0
1.5	0	0	0	0	0
2.0	0	0	0	0	0
2.5	0	0	0	0	0
3.0	0	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE B- 2

$$\sum = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

CASE 1-Bivariate

GILLIS

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

Sample Size

DIFFERENCE IN LOCATION OF EMITTERS	5	10	20	50	100
0.0	96	96	95	95	94
0.5	71	50	14	1	0
1.0	21	1	0	0	0
1.5	0	0	0	0	0
2.0	0	0	0	0	0
2.5	0	0	0	0	0
3.0	0	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Note that two sets of simulations were run for Sample Size 5.

Estimated probability of Type I Error = $1 - \frac{\text{Percentage}}{100}$

when difference in location is 0.

Estimated probability of Type II Error = $\frac{\text{Percentage}}{100}$

when difference in location is not 0.

Gillis Distribution, Case 2

The test assumes the underlying distribution is Normal, that the variance-covariance matrix is estimated from the data, and that the location is estimated by the mean of the data.

The first assumption is not met. The data follow a Gillis distribution, not the Normal distribution. However, the second and third assumptions are met.

Summary

Probability of Type I error:

Varies between 0.12 and 0.03.

Probability of Type II error:

Varies between 0.94 and 0 depending on sample size, assumed difference in location of emitters and magnitude of variance-covariance.

TABLE B-3

$$\sum = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

CASE 2

GILLIS

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

Sample Size

DIFFERENCE IN LOCATION OF EMITTERS	5	10	20	50	100
0.0	94	94	92	92	94
0.5	88	70	16	0	0
1.0	87	10	0	0	0
1.5	84	0	0	0	0
2.0	78	0	0	0	0
2.5	75	0	0	0	0
3.0	67	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE B-4

$$\sum = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

CASE 2

GILLIS

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size: 5	10	20	50	100
0.0	94	94	96	96	97
0.5	93	80	52	2	0
1.0	94	34	0	0	0
1.5	85	2	0	0	0
2.0	82	.05	0	0	0
2.5	82	0	0	0	0
3.0	74	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE B- 5

$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

CASE 2-Bivariate

GILLIS

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	88	96	98	97	97
0.5	86	67	30	0	0
1.0	75	4	0	0	0
1.5	48	0	0	0	0
2.0	22	0	0	0	0
2.5	4	0	0	0	0
3.0	0	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Note that two sets of simulations were run for Sample Size 5.

Estimated probability of Type I Error = $1- \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II Error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE B-6

$$\Sigma = \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix}$$

CASE 2-Bivariate

GILLIS

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size				
	5	10	20	50	100
0.0	90	94	92	92	96
0.5	88	84	71	11	0
1.0	86	42	4	0	0
1.5	69	4	0	0	0
2.0	49	0	0	0	0
2.5	22	0	0	0	0
3.0	5	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Note that two sets of simulations were run for Sample Size 5.

Estimated probability of Type I Error = $1 - \frac{\text{Percentage}}{100}$

when difference in location is 0.

Estimated probability of Type II Error = $\frac{\text{Percentage}}{100}$

when difference in location is not 0.

Gillis Distribution, Case 3

The test assumes the underlying distribution is Normal, that the variance-covariance matrix is known, and the location is estimated by the mean of the data.

The first two assumptions are not met. The data follow a Gillis distribution, not the Normal distribution, and the variance-covariance matrix used in the test is estimated from the data, not known. However, the third assumption is met.

Summary

Probability of Type I error:

Varies between 0.70 and 0.04 depending on sample size. Error rates are higher for 4-variate data.

Probability of Type II error:

Varies between 0.68 and 0. Error rates are higher for bivariate data.

TABLE B-7

$$\sum = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

CASE 3

GILLIS

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	30	70	87	94	96
0.5	15	17	4	0	0
1.0	1	0.5	0	0	0
1.5	0	0	0	0	0
2.0	0	0	0	0	0
2.5	0	0	0	0	0
3.0	0	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE B-8

$$\sum = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

CASE 3

GILLIS

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	30	70	89	90	92
0.5	17	40	22	1	0
1.0	8	2	0	0	0
1.5	0.5	0	0	0	0
2.0	0	0	0	0	0
2.5	0	0	0	0	0
3.0	0	0	0	0	0

Difference in location is the same for each "direction', i.e. .5 is in each direction.

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE B-9

$$\sum = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

CASE 3-Bivariate

GILLIS

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	79	84	92	94	94
0.5	58	48	18	0	0
1.0	20	19	2	0	0
1.5	0	0	0	0	0
2.0	0	0	0	0	0
2.5	0	0	0	0	0
3.0	0	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Note that two sets of simulations were run for Sample Size 5.

Estimated probability of Type I Error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II Error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE B-10

$$\sum = \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix}$$

CASE 3-Bivariate

GILLIS

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	78	88	89	96	90
0.5	65	68	48	8	0
1.0	39	18	2	0	0
1.5	12	0	0	0	0
2.0	0	0	0	0	0
2.5	0	0	0	0	0
3.0	0	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Note that two sets of simulations were run for Sample Size 5.

Estimated probability of Type I Error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II Error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

Gillis Distribution, Case 4

The test assumes the underlying distribution is Normal, that the variance-covariance matrix is known, and that the location is estimated by the mean of the data.

The first two assumptions are not met. The data follow the Gillis distribution, not the Normal distribution, and the variance-covariance matrix used in the test is varied from the true by a percentage (-4% to 4%). This could be thought of as an effect of calibration error. However, the third assumption is met.

Summary

Probability of Type I error:

Varies between 0.24 and 0 depending on the magnitude of the variance-covariance error.

Probability of Type II error:

Varies between 0.92 and 0 depending on sample size, assumed difference in location of emitters and magnitude of the variance-covariance error.

TABLE B-11

CASE 4

$$\sum = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

GILLIS

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

% Change in Covariance Matrix	Diff. in Loc. Of Emitter	Sample Size 5	10	20	50	100
-4%	0	74	76	78	70	80
	0.5	32	9	0	0	0
	1.0	0.5	0	0	0	0
	2.0	0	0	0	0	0
-3%	0	84	84	85	80	84
	0.5	42	16	2	0	0
	1.0	0.5	0	0	0	0
	2.0	0	0	0	0	0
-2%	0	91	94	90	89	89
	0.5	46	16	1	0	0
	1.0	1	0	0	0	0
	2.0	0	0	0	0	0
2%	0	100	99	98	87	98
	0.5	68	42	7	0	0
	1.0	4	0	0	0	0
	2.0	0	0	0	0	0
3%	0	98	100	98	100	99
	0.5	78	40	11	0	0
	1.0	8	0	0	0	0
	2.0	0	0	0	0	0
4%	0	100	99	100	100	100
	0.5	85	54	10	.05	0
	1.0	10	.05	0	0	0
	2.0	0	0	0	0	0

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$

For 0% change, see Table B-1 results.

TABLE B-12

CASE 4

$$\sum = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

GILLIS

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

% Change in Covariance Matrix	Diff. in Loc. Of Emitter	Sample Size 5	10	20	50	100
-4%	0	75	80	76	76	77
	0.5	47	22	7	0	0
	1.0	8	0.5	0	0	0
	2.0	0	0	0	0	0
-3%	0	80	87	84	82	84
	0.5	52	36	15	0	0
	1.0	10	0.5	0	0	0
	2.0	0	0	0	0	0
-2%	0	92	89	88	87	89
	0.5	69	45	10	0	0
	1.0	19	0.5	0	0	0
	2.0	0	0	0	0	0
2%	0	98	98	99	96	97
	0.5	89	73	34	3	0
	1.0	36	8	0	0	0
	2.0	0	0	0	0	0
3%	0	100	99	99	98	99
	0.5	91	75	40	1	0
	1.0	44	8	0.5	0	0
	2.0	0	0	0	0	0
4%	0	100	100	100	100	98
	0.5	92	78	50	4	0
	1.0	56	10	0.5	0	0
	2.0	0	0	0	0	0

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$

TABLE B-13

CASE 4-Bivariate

$$\sum = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

% Change in Covariance Matrix	Diff. in Loc. Of. Emitter	Sample Size 5	10	20	50	100
-4%	0	88	82	84	86	84
	0.5	48	42	8	0	0
	1.0	8	0	0	0	0
	2.0	0	0	0	0	0
-3%	0	91	86	86	88	85
	0.5	59	32	13	0	0
	1.0	7	0	0	0	0
	2.0	0	0	0	0	0
-2%	0	94	92	92	91	90
	0.5	65	43	18	0	0
	1.0	12	2	0	0	0
	2.0	0	0	0	0	0
2%	0	98	99	96	97	96
	0.5	80	51	26	2	0
	1.0	26	2	0	0	0
	2.0	0	0	0	0	0
3%	0	98	98	96	96	98
	0.5	78	68	32	0	0
	1.0	36	4	0	0	0
	2.0	0	0	0	0	0
4%	0	99	98	100	98	98
	0.5	82	66	29	0	0
	1.0	32	6	0	0	0
	2.0	0	0	0	0	0

$$\text{Estimated probability of Type I error} = 1 - \frac{\text{Percentage}}{100}$$

$$\text{Estimated probability of Type II error} = \frac{\text{Percentage}}{100}$$

APPENDIX C

Triangular-uniform Distribution

In the simulation results which follow, the test distribution is the Triangular-uniform distribution.

The Triangular-uniform is a medium fat-tailed distribution (see sketch). The four components (x_1,x_2,x_3,x_4), are generated essentially as follows:

$x_1 = u_2$ (uniform)

$x_2 = u_1 + u_2$ (triangular)

$x_3 = u_3$ (uniform)

$x_4 = u_3 + u_4$ (triangular)

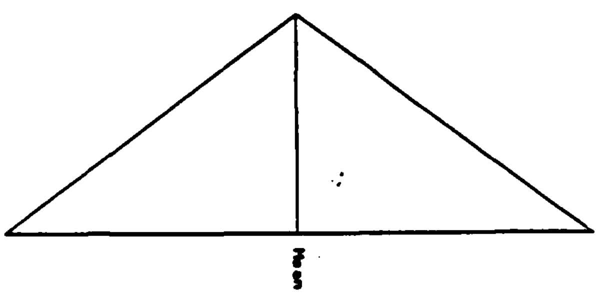
where u_1, u_2, u_3, and u_4 are independent uniform random variables.

x_1 and x_2 are dependent and x_3 and x_4 are dependent.

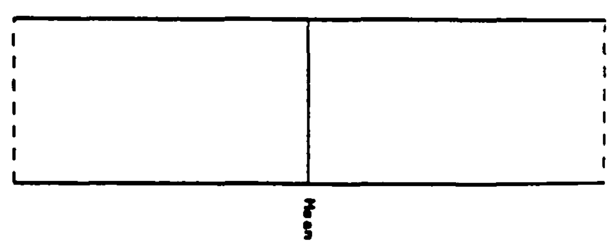
The formulas are identical for the bivariate simulations, except that only the first two components are used.

Sample sizes vary from 5 to 100. The tables which follow are categorized by the four basic assumptions about the data variance-covariance (see Page 14).

Triangular Distribution



Uniform Distribution



Triangular-uniform Distribution, Case 1

The test assumes the underlying distribution is Normal, that the variance-covariance matrix for the distribution is known, and that the location is estimated by the mean of the data.

The first assumption is not met. The data follow a Triangular-uniform distribution, not the Normal distribution. However, the second and third assumptions are met.

Summary

Probability of Type I error:

Agrees with the set level of 0.05.

Probability of Type II error:

Varies between 0.85 and 0 depending on sample size, assumed difference in location of emitters and magnitude of variance-covariance. Error is greater than when the distribution is Normal.

TABLE C-1

$$\sum = \begin{bmatrix} 1 & .7 & 0 & 0 \\ .7 & 1 & 0 & 0 \\ 0 & 0 & 1 & .7 \\ 0 & 0 & .7 & 1 \end{bmatrix}$$

CASE 1

TRIANGULAR-UNIFORM

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	96	96	96	96	95
0.5	76	45	30	0.5	0
1.0	22	2	0	0	0
1.5	2	0	0	0	0
2.0	0	0	0	0	0
2.5	0	0	0	0	0
3.0	0	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1- \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE C-2

$$\sum = \begin{bmatrix} 1 & .7 \\ .7 & 1 \end{bmatrix}$$

CASE 1-Bivariate

TRIANGULAR-UNIFORM

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	96	96	94	95	94
0.5	85	64	38	8	0
1.0	44	10	0	0	0
1.5	10	0	0	0	0
2.0	0	0	0	0	0
2.5	0	0	0	0	0
3.0	0	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Note that two sets of simulations were run for Sample Size 5.

Estimated probability of Type I Error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II Error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

Triangular-uniform Distribution, Case 2

The test assumes the underlying distribution is Normal, that the variance-covariance matrix is estimated from the data, and that the location is estimated by the mean of the data.

The first assumption is not met. The data follow a Triangular-uniform distribution, not the Normal distribution. However, the second and third assumptions are met.

Summary

Probability of Type I error:

Agrees with the set level of 0.05.

Probability of Type II error:

Varies between 0.95 and 0 depending on sample size, assumed difference in location of emitters and magnitude of variance-covariance. Error is greater than when the data follow the Normal distribution.

TABLE C-3

$$\sum = \begin{bmatrix} 1 & .7 & 0 & 0 \\ .7 & 1 & 0 & 0 \\ 0 & 0 & 1 & .7 \\ 0 & 0 & .7 & 1 \end{bmatrix}$$

CASE 2

TRIANGULAR-UNIFORM

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	95	92	96	94	95
0.5	95	78	38	2	0
1.0	89	35	0	0	0
1.5	88	2	0	0	0
2.0	83	0	0	0	0
2.5	-	-	-	-	-
3.0	-	-	-	-	-

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE C-4

$$\sum = \begin{bmatrix} 1 & .35 & 0 & 0 \\ .35 & .50 & 0 & 0 \\ 0 & 0 & .50 & .35 \\ 0 & 0 & .35 & 1 \end{bmatrix}$$

CASE 2

TRIANGULAR-UNIFORM

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size: 5	10	20	50	100
0.0	95	96	93	95	92
0.5	93	66	14	0	0
1.0	88	66	0	0	0
1.5	84	8	0	0	0
2.0	80	0	0	0	0
2.5	-	-	-	-	-
3.0	-	-	-	-	-

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1- \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE C-5

$$\sum = \begin{bmatrix} 1 & .7 \\ .7 & 1 \end{bmatrix}$$

CASE 2-Bivariate

TRIANGULAR-UNIFORM

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	93	96	94	96	96
0.5	90	78	50	5	0
1.0	78	32	2	0	0
1.5	59	2	0	0	0
2.0	34	0	0	0	0
2.5	15	0	0	0	0
3.0	4	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Note that two sets of simulations were run for Sample Size 5.

Estimated probability of Type I Error = $1- \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II Error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

Triangular-uniform Distribution, Case 3

The test assumes the underlying distribution is Normal, that the variance-covariance matrix is known, and the location is estimated by the mean of the data.

The first two assumptions are not met. The data follow a Triangular-uniform distribution, not the Normal distribution, and the variance-covariance matrix used in the test is estimated from the data, not known. However, the third assumption is met.

Summary

Probability of Type I error:

Varies between 0.75 and 0.05 depending mainly on sample size. Higher error rates occur for 4-variate data.

Probability of Type II error:

Varies between 0.65 and 0. Higher error rates occur for bivariate data.

TABLE C-6

$$\sum = \begin{bmatrix} 1 & .7 & 0 & 0 \\ .7 & 1 & 0 & 0 \\ 0 & 0 & 0 & .7 \\ 0 & 0 & .7 & 1 \end{bmatrix}$$

CASE 3

TRIANGULAR-UNIFORM

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size				
	5	10	20	50	100
0.0	25	67	84	93	95
0.5	16	36	18	0	0
1.0	5	0	0	0	0
1.5	.25	0	0	0	0
2.0	0	0	0	0	0
2.5	-	-	-	-	-
3.0	-	-	-	-	-

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE C-7

$$\sum = \begin{bmatrix} 1 & .35 & 0 & 0 \\ .35 & .5 & 0 & 0 \\ 0 & 0 & .5 & .35 \\ 0 & 0 & .35 & 1 \end{bmatrix}$$

CASE 3

TRIANGULAR-UNIFORM

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	30	74	88	93	94
0.5	11	21	1.5	0	0
1.0	.5	0	0	0	0
1.5	0	0	0	0	0
2.0	0	0	0	0	0
2.5	-	-	-	-	-
3.0	-	-	-	-	-

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error $= 1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error $= \frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE C-8

$$\sum = \begin{bmatrix} 1 & .7 \\ .7 & 1 \end{bmatrix}$$

CASE 3-Bivariate

TRIANGULAR-UNIFORM

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size				
	5	10	20	50	100
0.0	72	83	93	92	96
0.5	65	65	44	8	0
1.0	34	15	0	0	0
1.5	9	0	0	0	0
2.0	0	0	0	0	0
2.5	0	0	0	0	0
3.0	0	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Note that two sets of simulations were run for Sample Size 5.

Estimated probability of Type I Error = $1- \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II Error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

Triangular-uniform Distribution, Case 4

The test assumes the underlying distribution is Normal, that the variance-covariance matrix is known, and that the location is estimated by the mean of the data.

The first two assumptions are not met. The data follow the Triangular-uniform distribution, not the Normal distribution, and the variance-covariance matrix used in the test is varied from the true by a percentage (-4% to 4%). This could be thought of as an effect of calibration error. However, the third assumption is met.

Summary

Probability of Type I error:

Varies between 0.24 and 0 depending on the magnitude of the variance-covariance error.

Probability of Type II error:

Varies between 0.94 and 0 depending on sample size, assumed difference in location of emitters and magnitude of the variance-covariance error.

TABLE C-9

CASE 4

$$\sum = \begin{bmatrix} 1 & .7 & 0 & 0 \\ .7 & 1 & 0 & 0 \\ 0 & 0 & 1 & .7 \\ 0 & 0 & .7 & 1 \end{bmatrix}$$

TRIANGULAR-UNIFORM

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

% Change in Covariance Matrix	Diff. in Loc. Of Emitter	Sample Size 5	10	20	50	100
-4%	0	76	76	78	77	-
	0.5	48	22	6	0	-
	1.0	4	0	0	0	-
	2.0	0	0	0	0	-
-3%	0	82	85	86	88	-
	0.5	56	38	12	0	-
	1.0	7	.5	0	0	-
	2.0	0	0	0	0	-
-2%	0	94	90	88	86	-
	0.5	64	36	10	0	-
	1.0	11	1	0	0	-
	2.0	0	0	0	0	-
2%	0	98	97	98	98	-
	0.5	88	66	33	0	-
	1.0	34	3	0	0	-
	2.0	0	0	0	0	-
3%	0	99	98	99	98	-
	0.5	86	72	36	0	-
	1.0	30	7	0	0	-
	2.0	0	0	0	0	-
4%	0	99	98	98	100	-
	0.5	91	75	36	1	-
	1.0	42	6	0	0	-
	2.0	0	0	0	0	-

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$

For 0% change see Table C-1.

TABLE C-10

CASE 4-Bivariate

$$\sum = \begin{bmatrix} 1 & .7 \\ .7 & 1 \end{bmatrix}$$

TRIANGULAR-UNIFORM

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

% Change in Covariance Matrix: % Change	Diff. in Loc. Of. Emitter	Sample Size: 5	10	20	50	100
-4%	0	86	84	83	86	85
	0.5	67	42	22	4	0
	1.0	24	6	0	0	0
	2.0	0	0	0	0	0
-3%	0	86	88	88	88	84
	0.5	71	49	24	2	0
	1.0	32	9	0	0	0
	2.0	0	0	0	0	0
-2%	0	92	90	94	92	92
	0.5	75	60	30	4	0
	1.0	37	7	0	0	0
	2.0	0	0	0	0	0
2%	0	97	100	97	98	98
	0.5	86	76	50	10	0
	1.0	52	20	0	0	0
	2.0	2	0	0	0	0
3%	0	99	98	100	98	98
	0.5	94	82	52	11	0
	1.0	60	22	2	0	0
	2.0	1	0	0	0	0
4%	0	100	99	98	98	97
	0.5	92	78	60	16	0
	1.0	61	29	2	0	0
	2.0	1	0	0	0	0

$$\text{Estimated probability of Type I error} = 1 - \frac{\text{Percentage}}{100}$$

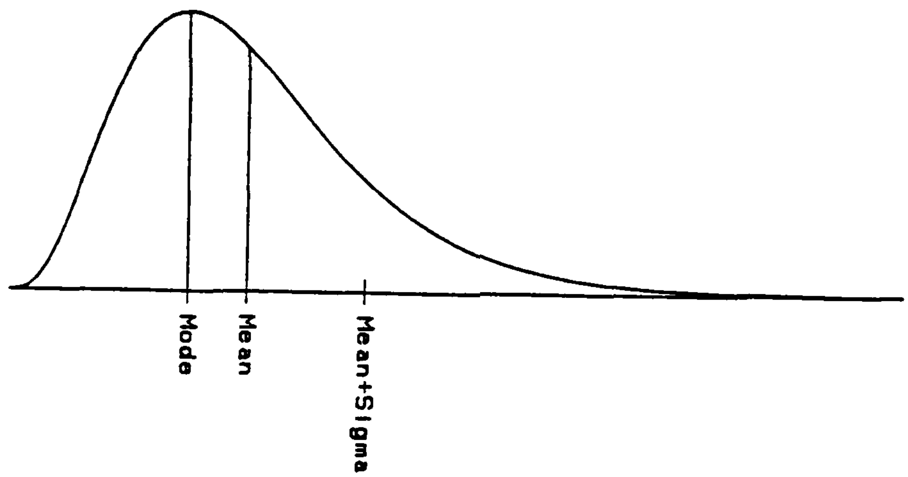
$$\text{Estimated probability of Type II error} = \frac{\text{Percentage}}{100}$$

APPENDIX D

Gamma Distribution

In the simulation results which follow, the test distribution is the Gamma distribution.

The Gamma distribution is a skewed distribution whose mean (expected) value is to the right of the mode (most frequent) value.



The density sketched above follows the general equation

$$f(x) = \frac{\lambda^n x^{n-1} e^{-\lambda x}}{\Gamma(n)}$$

Three types of this density were simulated in the 4-variate case:

a) n=2, coefficient of skewness = 1.414.
b) n=4, coefficient of skewness = 1.000.
c) n=9, coefficient of skewness = 0.666.

(Note that the coefficient of skewness for symmetric distributions such as Normal, Gillis, or Triangular-uniform is zero). As n increases the Gamma distribution approaches the Normal distribution.

The data are vectors with four components (x_1, x_2, x_3, x_4). These components are generated as follows:

$$x_1 = a_1 y_1$$

$$x_2 = a_2 y_1 + a_3 y_2$$

$$x_3 = a_4 y_3$$

$$x_4 = a_5 y_3 + a_6 y_4$$

where y_1, y_2, y_3 and y_4 are independent gamma variables and the a_i's are constants. (For each simulation case the y's were generated with the same n value but with varying lambda values). For bivariate cases, only the first two components were used.

Sample sizes vary from 5 to 100. The tables which follow are categorized by the four basic assumptions about the variance-covariance (see Page 14).

Gamma Distribution, Case 1

The test assumes the underlying distribution is Normal, that the variance-covariance matrix for the distribution is known, and that the location is estimated by the mean of the data.

The first assumption is not met. The data follow a Gamma distribution, not the Normal distribution. However, the second and third assumptions are met.

Summary

Probability of Type I error:

Agrees with the set level of 0.05.

Probability of Type II error:

Varies between 0.76 and 0 depending on sample size, assumed difference in location of emitters and magnitude of variance-covariance. Error is greater than when the distribution is Normal.

TABLE D-1

$$\sum = \begin{bmatrix} 1 & .7 & 0 & 0 \\ .7 & 1 & 0 & 0 \\ 0 & 0 & 1 & .7 \\ 0 & 0 & .7 & 1 \end{bmatrix}$$

CASE 1

GAMMA (n = 2)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size: 5	10	20	50	100
0.0	92	94	94	92	94
0.5	76	54	18	0.5	0
1.0	19	2	0	0	0
1.5	3	0	0	0	0
2.0	0	0	0	0	0
2.5	0	0	0	0	0
3.0	0	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE D-2

$$\sum = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

CASE 1-Bivariate

GAMMA (n=3)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size				
	5	10	20	50	100
0.0	97	94	96	97	96
0.5	67	56	14	2	0
1.0	15	2	0	0	0
1.5	0	0	0	0	0
2.0	0	0	0	0	0
2.5	0	0	0	0	0
3.0	0	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Note that two sets of simulations were run for Sample Size 5.

Estimated probability of Type I Error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II Error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

Gamma Distribution, Case 2

The test assumes the underlying distribution is Normal, that the variance-covariance matrix is estimated from the data, and that the location is estimated by the mean of the data.

The first assumption is not met. The data follow a Gamma distribution, not the Normal distribution. However, the second and third assumptions are met.

Summary

Probability of Type I error:

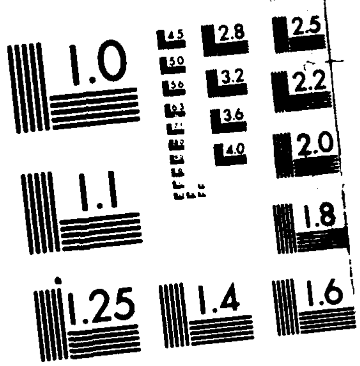
Varies from 0.12 to 0.03. Small sample sizes for bivariate data have the highest error rate.

Probability of Type II error:

Varies between 0.92 and 0 depending mainly on sample size, assumed difference in location of emitters and magnitude of variance-covariance. Error is greater than when the data follow the Normal distribution.

AD-A166 477 POWER OF STATISTICAL TESTS USED IN CORRELATION TECHNIQUES FOR BATTLEFIELD IDENTIFICATION(U) JET PROPULSION LAB PASADENA CA J MYHRE ET AL. AUG 85 2/2

UNCLASSIFIED JPL-D-2793 NAS7-918 F/G 12/1 NL



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS 1963 A

TABLE D-3

$$\sum = \begin{bmatrix} 4 & 1 & 0 & 0 \\ 1 & 4 & 0 & 0 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 1 & 4 \end{bmatrix}$$

CASE 2

GAMMA (n = 2)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	94	90	91	95	95
0.5	92	72	58	22	7.5
1.0	90	49	16	0	0
1.5	85	24	2	0	0
2.0	85	6.5	0	0	0
3.0	78	0	0	0	0
4.0	72	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE D-4

$$\sum = \begin{bmatrix} 1 & .7 & 0 & 0 \\ .7 & 1 & 0 & 0 \\ 0 & 0 & 1 & .7 \\ 0 & 0 & .7 & 1 \end{bmatrix}$$

CASE 2

GAMMA (n = 4)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size				
	5	10	20	50	100
0.0	95	92	92	92	95
0.5	92	64	34	4	0
1.0	90	18	0	0	0
1.5	86	1	0	0	0
2.0	78	0	0	0	0
2.5	-	-	-	-	-
3.0	-	-	-	-	-

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1- \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE D-5

$$\sum = \begin{bmatrix} 1 & .7 & 0 & 0 \\ .7 & 1 & 0 & 0 \\ 0 & 0 & 1 & .7 \\ 0 & 0 & .7 & 0 \end{bmatrix}$$

CASE 2

GAMMA (n = 9)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size				
	5	10	20	50	100
0.0	97	93	97	92	-
0.5	90	60	39	5	0
1.0	89	16	0	0	0
1.5	85	4	0	0	0
2.0	80	0	0	0	0
2.5	-	-	-	-	-
3.0	-	-	-	-	-

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE D-6

$$\sum = \begin{bmatrix} 1 & .35 & 0 & 0 \\ .35 & .50 & 0 & 0 \\ 0 & 0 & .5 & .35 \\ 0 & 0 & .35 & .5 \end{bmatrix}$$

CASE 2

GAMMA (n = 9)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size				
	5	10	20	50	100
0.0	93	94	93	96	96
0.5	90	42	6	0	0
1.0	83	3.5	0	0	0
1.5	77	0	0	0	0
2.0	74	0	0	0	0
2.5	-	-	-	-	-
3.0	-	-	-	-	-

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE D-7

$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

CASE 2-Bivariate

GAMMA (n=3)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	88	88	96	96	94
0.5	76	52	16	2	0
1.0	49	7	0	0	0
1.5	27	0	0	0	0
2.0	12	0	0	0	0
2.5	4	0	0	0	0
3.0	1	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Note that two sets of simulations were run for Sample Size 5.

Estimated probability of Type I Error = $1 - \frac{\text{Percentage}}{100}$

when difference in location is 0.

Estimated probability of Type II Error = $\frac{\text{Percentage}}{100}$

when difference in location is not 0.

TABLE D-8

$$\Sigma = \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix}$$

CASE 2-Bivariate

GAMMA (n=2)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	90	88	92	93	94
0.5	72	33	2	0	0
1.0	48	14	1	0	0
1.5	26	2	0	0	0
2.0	12	0	0	0	0
2.5	5	0	0	0	0
3.0	2	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Note that two sets of simulations were run for Sample Size 5.

Estimated probability of Type I Error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II Error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

Gamma Distribution, Case 3

The test assumes the underlying distribution is Normal, that the variance-covariance matrix is known, and the location is estimated by the mean of the data.

The first two assumptions are not met. The data follow a Gamma distribution, not the Normal distribution, and the variance-covariance matrix used in the test is estimated from the data, not known. However, the third assumption is met.

Summary

Probability of Type I error:

Varies between 0.80 and 0.06 depending on sample size. Higher error rates occur for 4-variate data.

Probability of Type II error:

Varies between 0.42 and 0 depending mainly on sample size, assumed difference in location of emitters and magnitude of variance-covariance.

TABLE D-9

$$\sum = \begin{bmatrix} 4 & 1 & 0 & 0 \\ 1 & 4 & 0 & 0 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 1 & 4 \end{bmatrix}$$

CASE 3

GAMMA (n = 2)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size: 5	10	20	50	100
0.0	20	64	82	88	94
0.5	13	39	40	22	6
1.0	6	17	9	0	0
1.5	4	3	1	0	0
2.0	1	0	0	0	0
3.0	0	0	0	0	0
4.0	0	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE D-10

$$\sum = \begin{bmatrix} 1 & .7 & 0 & 0 \\ .7 & 1 & 0 & 0 \\ 0 & 0 & 1 & .7 \\ 0 & 0 & .7 & 1 \end{bmatrix}$$

CASE 3

GAMMA (n = 4)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size				
	5	10	20	50	100
0.0	27	67	84	91	90
0.5	12	20	12	2	0
1.0	3	2	0	0	0
1.5	0	0	0	0	0
2.0	0	0	0	0	0
2.5	-	-	-	-	-
3.0	-	-	-	-	-

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE D-11

$$\sum = \begin{bmatrix} 1 & .7 & 0 & 0 \\ .7 & 1 & 0 & 0 \\ 0 & 0 & 1 & .7 \\ 0 & 0 & .7 & 1 \end{bmatrix}$$

CASE 3

GAMMA (n = 9)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size				
	5	10	20	50	100
0.0	24	64	86	91	85
0.5	14	28	18	4	.5
1.0	5	2.5	0	0	0
1.5	.75	0	0	0	0
2.0	.25	0	0	0	0
2.5	-	-	-	-	-
3.0	-	-	-	-	-

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE D-12

$$\Sigma = \begin{bmatrix} 1 & .35 & 0 & 0 \\ .35 & .50 & 0 & 0 \\ 0 & 0 & .50 & .35 \\ 0 & 0 & .35 & 1 \end{bmatrix}$$

CASE 3

GAMMA (n = 4)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	26	62	82	91	92
0.5	4	16	3	0	0
1.0	.5	0	0	0	0
1.5	0	0	0	0	0
2.0	0	0	0	0	0
2.5	-	-	-	-	-
3.0	-	-	-	-	-

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE D-13

$$\sum = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

CASE 3-Bivariate

GAMMA (n=3)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	65	87	90	92	90
0.5	41	29	20	0	0
1.0	14	4	0	0	0
1.5	4	0	0	0	0
2.0	0	0	0	0	0
2.5	0	0	0	0	0
3.0	0	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Note that two sets of simulations were run for Sample Size 5.

Estimated probability of Type I Error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II Error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

TABLE D-14

$$\Sigma = \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix}$$

CASE 3-Bivariate

GAMMA (n=2)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	72	86	89	90	94
0.5	42	38	20	1	0
1.0	14	4	0	0	0
1.5	3	0	0	0	0
2.0	2	0	0	0	0
2.5	0	0	0	0	0
3.0	0	0	0	0	0

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Note that two sets of simulations were run for Sample Size 5.

Estimated probability of Type I Error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II Error = $\frac{\text{Percentage}}{100}$ when difference in location is not 0.

Gamma Distribution, Case 4

The test assumes the underlying distribution is Normal, that the variance-covariance matrix is known, and that the location is estimated by the mean of the data.

The first two assumptions are not met. The data follow the Gamma distribution, not the Normal distribution, and the variance-covariance matrix used in the test is varied from the true by a percentage (-4% to 4%). This could be thought of as an effect of calibration error. However, the third assumption is met.

Summary

Probability of Type I error:

Varies between 0.24 and 0.02 depending on the magnitude of the variance-covariance error.

Probability of Type II error:

Varies between 0.96 and 0 depending on sample size, assumed difference in location of emitters and magnitude of the variance-covariance error.

TABLE D-15

CASE 4

$$\sum = \begin{bmatrix} 1 & .7 & 0 & 0 \\ .7 & 1 & 0 & 0 \\ 0 & 0 & 1 & .7 \\ 0 & 0 & .7 & 1 \end{bmatrix}$$

GAMMA (n = 2)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

% Change in Covariance Matrix	Diff. in Loc. Of Emitter	Sample Size 5	10	20	50	100
-4%	0	79	82	76	79	-
	0.5	47	23	2	0	-
	1.0	3	0	0	0	-
	2.0	0	0	0	0	-
-3%	0	84	80	82	78	-
	0.5	46	26	8	0	-
	1.0	7	0	0	0	-
	2.0	0	0	0	0	-
-2%	0	91	88	90	91	-
	0.5	58	40	14	1	-
	1.0	6	1	0	0	-
	2.0	0	0	0	0	-
2%	0	96	98	98	97	-
	0.5	88	70	24	0	-
	1.0	28	4	0	0	-
	2.0	0	0	0	0	-
3%	0	98	97	98	96	-
	0.5	86	74	40	1	-
	1.0	38	4	0	0	-
	2.0	0	0	0	0	-
4%	0	97	98	98	98	-
	0.5	94	78	42	4	-
	1.0	36	6	0	0	-
	2.0	0	0	0	0	-

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$

For 0% Change see Table D-1 results.

TABLE D-16

CASE 4-Bivariate

$$\sum = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

GAMMA (n=3)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

% Change in Covariance Matrix	Diff. in Loc. Of Emitter	Sample Size 5	10	20	50	100
-4%	0	90	88	87	86	82
	0.5	44	31	6	0	0
	1.0	8	0	0	0	0
	2.0	0	0	0	0	0
-3%	0	81	83	84	86	86
	0.5	48	34	15	0	0
	1.0	7	1	0	0	0
	2.0	0	0	0	0	0
-2%	0	90	92	93	92	90
	0.5	60	40	10	1	0
	1.0	4	0	0	0	0
	2.0	0	0	0	0	0
2%	0	94	98	93	98	94
	0.5	82	60	22	0	0
	1.0	20	2	0	0	0
	2.0	0	0	0	0	0
3%	0	95	99	96	98	98
	0.5	88	70	34	2	0
	1.0	26	8	0	0	0
	2.0	0	0	0	0	0
4%	0	99	100	98	100	98
	0.5	96	70	26	1	0
	1.0	24	4	0	0	0
	2.0	0	0	0	0	0

Estimated probability of Type I error $= 1 - \frac{\text{Percentage}}{100}$

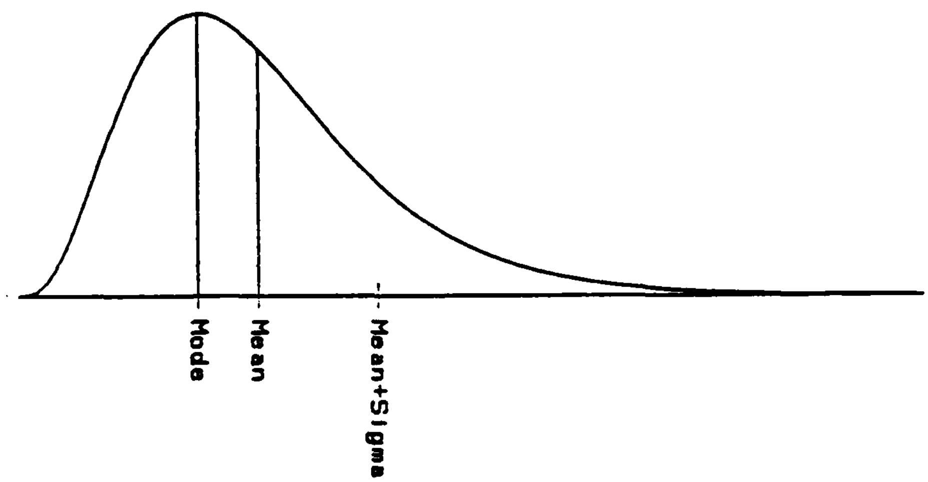
Estimated probability of Type II error $= \frac{\text{Percentage}}{100}$

APPENDIX E

Gamma Mode Distribution

In the simulation results which follow the test distribution is the Gamma distribution.

The Gamma distribution is a skewed distribution whose mean (expected) value is to the right of the mode (most frequent) value. For the simulation cases which follow the location was estimated by the mode (most frequent) value rather than the mean (expected) value of the data.



The density sketched above follows the general equation

$$f(x) = \frac{\lambda^n x^{n-1} e^{-\lambda x}}{\Gamma(n)}$$

The data are vectors with four components (x_1, x_2, x_3, x_4). These components are generated as follows:

$$x_1 = a_1 y_1$$

$$x_2 = a_2 y_1 + a_3 y_2$$

$$x_3 = a_4 y_3$$

$$x_4 = a_5 y_3 + a_6 y_4$$

where y_1, y_2, y_3 and y_4 are independent gamma variables and the a_i's are constants. (For each simulation case the y's were generated with the same n value but with varying lambda values). For the bivariate case, only the first two components were used.

Sample sizes vary from 5 to 100. The tables which follow are categorized by the four basic assumptions about the variance-covariance (see Page 14).

Gamma Mode Distribution, Case 1

The test assumes the underlying distribution is Normal, that the variance-covariance matrix for the distribution is known, and that the location is estimated by the mean of the data.

The first and third assumptions are not met. The data follow a Gamma distribution, not the Normal distribution. Also, the location is estimated by the mode of the data. However, the second assumption is met.

Summary

Probability of Type I error:

Varies between 1.0 and 0.88 depending on sample size. The larger the sample size the larger the probability of error.

Probability of Type II error:

Almost 0.

TABLE E-1

$$\Sigma = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

CASE 1

GAMMA (n = 2)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size				
	5	10	20	50	100
0.0	12	4	0	0	0
0.5	2	0	0	0	0
1.0	0	0	0	0	0
1.5	0	0	0	0	0
2.0	0	0	0	0	0
2.5	0	0	0	0	0
3.0	0	0	0	0	0

Mean: (1.414, 1.414, 1.414, 1.414)
Mode: (0.707, 0.707, 0.707, 0.707)

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$

TABLE E-2

$$\sum = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

CASE 1-Bivariate

GAMMA (n=3)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	42	22	18	0	0
0.5	16	2	0	0	0
1.0	1	0	0	0	0
1.5	0	0	0	0	0
2.0	0	0	0	0	0
2.5	0	0	0	0	0
3.0	0	0	0	0	0

Mean: (1.732, 1.732)

Mode: (1.155, 1.155)

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1-\frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$

Gamma Mode Distribution, Case 2

The test assumes the underlying distribution is Normal, that the variance-covariance matrix is estimated from the data, and that the location is estimated by the mean of the data.

The first and third assumptions are not met. The data follow a Gamma distribution, not the Normal distribution. Also, the location is estimated by the mode of the data. However, the second assumption is met.

Summary

Probability of Type I error:

Varies between 1.0 and 0.09 depending on sample size. The larger the sample size the larger the probability of error.

Probability of Type II error:

Varies between 0.88 for small sample size and 0 for larger sample size. Also depends on assumed difference in location of emitters. Error is less than when the distribution is Normal.

TABLE E-3

$$\sum = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

CASE 2

GAMMA (n = 2)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	91	37	0.4	0	0
0.5	86	1	0	0	0
1.0	79	0	0	0	0
1.5	73	0	0	0	0
2.0	64	0	0	0	0
2.5	59	0	0	0	0
3.0	55	0	0	0	0

Mean: (1.414, 1.414, 1.414, 1.414)
Mode: (0.707, 0.707, 0.707, 0.707)

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$

TABLE E-4

$$\sum = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

CASE 2

GAMMA (n = 3)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	92	42	5	0	0
0.5	88	7	0	0	0
1.0	87	0.5	0	0	0
1.5	78	0	0	0	0
2.0	74	0	0	0	0
2.5	73	0	0	0	0
3.0	65	0	0	0	0

Mean: (1.732, 2.449, 3.000, 3.464)
Mode: (1.155, 1.633, 2.000, 2.309)

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$

TABLE E-5

$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

CASE 2-Bivariate

GAMMA (n=3)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	86	48	24	0	0
0.5	52	7	0	0	0
1.0	27	0	0	0	0
1.5	6	0	0	0	0
2.0	4	0	0	0	0
2.5	1	0	0	0	0
3.0	0	0	0	0	0

Mean: (1.732, 1.732)
Mode: (1.155, 1.155)

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$

Gamma Mode Distribution, Case 3

The test assumes the underlying distribution is Normal, that the variance-covariance matrix is known, and the location is estimated by the mean of the data.

None of the assumptions are met. The data follow a Gamma distribution, not the Normal distribution. The variance-covariance matrix used in the test is estimated from the data, not known. The location is estimated by the mode of the data.

Summary

Probability of Type I error:

Varies between 1.0 and 0.93 depending on sample size. The larger the sample size the larger the probability of error.

Probability of Type II error:

Almost 0.

TABLE E-6

$$\sum = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

CASE 3

GAMMA (n = 2)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size				
	5	10	20	50	100
0.0	4	3	0.4	0	0
0.5	0.1	0	0	0	0
1.0	0	0	0	0	0
1.5	0	0	0	0	0
2.0	0	0	0	0	0
2.5	0	0	0	0	0
3.0	0	0	0	0	0

Mean: (1.414, 1.414, 1.414, 1.414)
Mode: (0.707, 0.707, 0.707, 0.707)

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$

TABLE E-7

$$\Sigma = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

CASE 3

GAMMA (n = 3)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size				
	5	10	20	50	100
0.0	7	6	0.5	0	0
0.5	1	0	0	0	0
1.0	0.5	0	0	0	0
1.5	0	0	0	0	0
2.0	0	0	0	0	0
2.5	0	0	0	0	0
3.0	0	0	0	0	0

Mean: (1.732, 2.449, 3.000, 3.464)
Mode: (1.155, 1.633, 2.000, 2.309)

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1- \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$

TABLE E-8

$$\sum = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

CASE 3-Bivariate

GAMMA (n=2)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

DIFFERENCE IN LOCATION OF EMITTERS	Sample Size 5	10	20	50	100
0.0	44	32	14	0	0
0.5	9	2	0	0	0
1.0	1	0	0	0	0
1.5	0	0	0	0	0
2.0	0	0	0	0	0
2.5	0	0	0	0	0
3.0	0	0	0	0	0

Mean: (1.732, 1.732)
Mode: (1.155, 1.155)

Difference in location is the same for each "direction", i.e. .5 is in each direction.

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$ when difference in location is 0.

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$

Gamma Mode Distribution, Case 4

The test assumes the underlying distribution is Normal, that the variance-covariance matrix is known, and that the location is estimated by the mean of the data.

None of the assumptions are met. The data follow the Gamma distribution, not the Normal distribution, and the variance-covariance matrix used in the test is varied from the true by a percentage (-4% to 4%). This could be thought of as an effect of calibration error. The location is estimated by the mode of the data.

Summary

Probability of Type I error:

Varies between 1.0 and 0.76 depending on sample size and the magnitude of the variance-covariance error. The larger the sample size the larger the probability of type I error.

Probability of Type II error:

From 0.24 to 0.

TABLE E-9

CASE 4

$$\sum = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

GAMMA (n = 2)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

% Change in Covariance Matrix	Diff. in Loc. Of Emitter	Sample Size 5	10	20	50	100
-4%	0	24	10	2	0	0
	0.5	1	0	0	0	0
	1.0	0	0	0	0	0
	2.0	0	0	0	0	0
-3%	0	20	7	4	0	0
	0.5	2	0	0	0	0
	1.0	0	0	0	0	0
	2.0	0	0	0	0	0
-2%	0	24	6	0.5	0	0
	0.5	0.5	0	0	0	0
	1.0	0	0	0	0	0
	2.0	0	0	0	0	0
2%	0	18	2	0	0	0
	0.5	2	0.5	0	0	0
	1.0	0	0	0	0	0
	2.0	0	0	0	0	0
3%	0	13	2	0	0	0
	0.5	0.5	0	0	0	0
	1.0	0	0	0	0	0
	2.0	0	0	0	0	0
4%	0	13	0.5	0	0	0
	0.5	2	0	0	0	0
	1.0	0	0	0	0	0
	2.0	0	0	0	0	0

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$

For 0% change see Table E-1.

Mean: (1.414, 1.414, 1.414, 1.414)
Mode: (0.707, 0.707, 0.707, 0.707)

TABLE E-10

CASE 4

$$\sum = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

GAMMA (n = 3)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

% Change in Covariance Matrix	Diff. in Loc. Of Emitter	Sample Size 5	10	20	50	100
-4%	0	33	15	6	0.5	0
	0.5	6	0	0	0	0
	1.0	2	0	0	0	0
	2.0	0	0	0	0	0
-3%	0	45	22	8	0	0
	0.5	6	0.5	0	0	0
	1.0	0.5	0	0	0	0
	2.0	0	0	0	0	0
-2%	0	34	15	2	0.5	0
	0.5	8	0.5	0	0	0
	1.0	1	0	0	0	0
	2.0	0	0	0	0	0
2%	0	24	4	0.5	0	0
	0.5	3	0	0	0	0
	1.0	0.5	0	0	0	0
	2.0	0	0	0	0	0
3%	0	16	5	0	0	0
	0.5	6	0.5	0	0	0
	1.0	0.5	0	0	0	0
	2.0	0	0	0	0	0
4%	0	21	4	0.5	0	0
	0.5	2	0.5	0	0	0
	1.0	0	0	0	0	0
	2.0	0	0	0	0	0

Estimated probability of Type I error = $1 - \frac{\text{Percentage}}{100}$

Estimated probability of Type II error = $\frac{\text{Percentage}}{100}$

Mean: (1.732, 2.449, 3.000, 3.464)
Mode: (1.155, 1.633, 2.000, 2.309)

TABLE E-11

CASE 4-Bivariate

$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

GAMMA (n=3)

PERCENTAGE OF LOCATIONS ACCEPTED AS THE SAME

% Change in Covariance Matrix: % Change	Diff. in Loc. Of Emitter	Sample Size: 5	10	20	50	100
-4%	0	46	50	26	8	0
	0.5	24	2	0	0	0
	1.0	3	0	0	0	0
	2.0	0	0	0	0	0
-3%	0	50	47	16	4	0
	0.5	18	3	0	0	0
	1.0	0	0	0	0	0
	2.0	0	0	0	0	0
-2%	0	55	36	16	2	0
	0.5	12	2	0	0	0
	1.0	0	0	0	0	0
	2.0	0	0	0	0	0
2%	0	49	20	2	0	0
	0.5	10	0	0	0	0
	1.0	0	0	0	0	0
	2.0	0	0	0	0	0
3%	0	45	26	6	0	0
	0.5	13	2	0	0	0
	1.0	1	0	0	0	0
	2.0	0	0	0	0	0
4%	0	44	18	2	0	0
	0.5	6	1	0	0	0
	1.0	1	0	0	0	0
	2.0	0	0	0	0	0

Estimated probability of Type I error $= 1 - \frac{\text{Percentage}}{100}$

Estimated probability of Type II error $= \frac{\text{Percentage}}{100}$

APPENDIX F

CASE 1

Mathematical Tests

H_0: $\underline{\mu} = \underline{\mu}_0$ fixed, Σ known

H_1: $\underline{\mu} \neq \underline{\mu}_0$

The data are $\underline{x}_i = (x_{i1}, x_{i2}, \ldots, x_{ip})$, $i = 1, \ldots, n$.

For this report p = 2 or 4. The data are assumed to be independent observations for a p-variate (4-variate) normal distribution with mean $\underline{\mu} = (\mu_1, \mu_2, \ldots, \mu_p)$ and the variance-covariance matrix

$$\sum = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_p^2 \end{bmatrix}$$

Let α be the level of significance of the test that is, the probability we reject H_0 when it is true.

CASE 1

We estimate $\underline{\mu}$ by

$$\underline{\bar{x}} = \frac{1}{n} \left[\sum_{i=1}^{n} x_{i1} \quad \sum_{i}^{n} x_{i2}, \; \ldots, \; \sum_{i}^{n} x_{ip} \right]$$

The test is to reject H_0 if $T^2 = n\,(\underline{\bar{x}} - \underline{\mu}_0)\Sigma^{-1}(\underline{\bar{x}} - \underline{\mu}_0)^T \geq \chi_p^2\,(\alpha)$ where $\chi_p^2\,(\alpha)$ is the 100 $(1-\alpha)$ percentile point of a chi-square disbribution with p degrees of freedom*.

When $\underline{\mu} \neq \underline{\mu}_0$, T^2 is distributed as a non-central chi-square with p degrees of freedom and non-centrality parameter

$$n\,(\underline{\mu} - \underline{\mu}_0)\quad \Sigma^{-1}\,(\underline{\mu} - \underline{\mu}_0)^T.$$

---

*In this case the test statistic T^2 is called Hotellings T^2 statistic.

CASE 2

H_0: $\underline{\mu} = \underline{\mu}_0$ fixed, Σ unknown

H_1, $\underline{\mu} \neq \underline{\mu}_0$

The data follow the same assumptions as for Case 1 except that Σ is unknown. We estimate Σ by the matrix

$$S = \frac{1}{n-1} \sum_{i=1}^{n} (\underline{x}_i - \bar{\underline{x}})(\underline{x}_i - \bar{\underline{x}})^T$$

The test is to reject H_0 if $T^2 = n\,(\bar{\underline{x}} - \underline{\mu}_0)\ S^{-1}\,(\bar{\underline{x}} - \underline{\mu}_0)^T \geq T_0{}^2$

where $T_0{}^2 = \frac{(n-1)}{n-p} p \quad F_{p,n-p}\ (\alpha)$.

$F_{p,n-p}(\alpha)$ is the 100 $(1-\alpha)$ percentile point of an F distribution with p,n-p degrees of freedom.

When $\underline{\mu} = \underline{\mu}_0$, T^2 is distributed as non-central F distribution.

For details see Anderson, T.W., An Introduction to Multivariate Statistical Analysis. New York: John Wiley & Sons, Inc. 1958, Pages 113-115.

CASE 3

H_0: $\underline{\mu} = \underline{\mu}_0$ fixed, Assume $\Sigma = S$

H_1: $\underline{\mu} \neq \underline{\mu}_0$

The data follow the same assumptions as for Case 1 except that we assume our estimate of Σ, S, is in fact equal to Σ.

The test decision is to reject H_0 if

$$T^2 = n\ (\underline{\bar{x}} - \underline{\mu}_0)\ \ S^{-1}\ (\underline{\bar{x}} - \underline{\mu}_0)^T \geq \chi_p^2\ (\alpha)$$

where χ_p^2 (α) is the 100 (1-α) percentile point of a chi-square distribution with p degrees of freedom. α is the level of significance (probability of rejection H_0 when it is true).

In this case the assumption that $\Sigma = S$ will not be valid for small sample sizes and the probability of rejecting H_0 when it is true will be much greater than α.

CASE 4

$H_0: \underline{\mu} = \underline{\mu}_0$ fixed, Σ assumed known

$H_1: \underline{\mu} \neq \underline{\mu}_0$

The data follow the same underlying assumptions as for Case 1 except that our assumed known Σ is *incorrect*. The correct variance-covariance matrix is $\alpha\Sigma$.

The test is to reject H_0 if

$$T^2 = n\ (\underline{\bar{x}} - \underline{\mu}_0)\ \Sigma^{-1}\ (\underline{\bar{x}} - \underline{\mu}_0)^T \geq \chi^2_p\ (\alpha)$$

where $\chi^2_p\ (\alpha)$ is the 100 $(1-\alpha)$ percentile point of a chi-square distribution with p degrees of freedom. α is the level of significance (probability of rejecting H_0 when it is true).

In this case the T^2 is incorrect. If we knew "α" it would be correct to use the T^2 below as the test statistic. Reject if

$$T^2 = n\ (\underline{\bar{x}} - \underline{\mu}_0\ )\ (\alpha\Sigma)^{-1}\ (\underline{\bar{x}} - \underline{\mu}_0)^T \geq \chi^2_p\ (\alpha)$$

APPENDIX G

Simulation Procedures

The test error frequency values cited in the previous appendices were generated through a simulation technique. More explicitly, a VAX-11/780 mini-computer located at Claremont McKenna College was used to generate observations from various distributions (Normal, Gillis, etc.), and the tests were repeated 200 to 500 times. The figures listed in the tables are the percentage of the time the tests accepted. For example, if for a given case the test accepted 190 out of 200 times, the reported figure would be 95%.

The simulations were conducted in the following manner. First, a case was chosen:

1) The location and parametrics of the emitter in the database were specified.

2) The distribution of the incoming data was specified (Normal, Gillis, etc.).

3) The test to be used was specified, corresponding to the four cases listed on page 14.

4) The parameters of the data distribution were specified; for example, the mean and variance-covariance for the Normal. In order to determine the probability of type II error, the data were frequently generated around a mean point <u>other</u> than the true location of the emitter.

5) The sample size (number of data observations) was specified.

Then, for this particular case, the following steps were performed 200 to 500 times:

1) Generate n observations from the given distribution (where n is the sample size).

2) Conduct the chosen test.

3) Record acceptance or rejection.

APPENDIX H

RELATING ELLIPSE SHAPE WITH MEASUREMENT ERROR

This appendix explains how the shape of a confidence ellipse is related to the associated covariance matrix. Note that the geometrical analysis of this appendix applies primarily to the bivariate simulations.

The covariance matrices Σ used in this report measure the error in the individual observations. The covariance matrix which reflects the combination of this data is Σ/n, where n is sample size. The equation of the 95% confidence ellipse is

$$(\underline{x} - \underline{\mu})^T(\Sigma/n)^{-1}(\underline{x} - \underline{\mu}) = \chi^2_2(.95) = 5.991$$

Increased sample size, n, decreases the amount of measurement error. The ultimate effect of this is to decrease the size of the confidence ellipse. When sample size is not mentioned in this report assume that n=1.

Let A be the variance in the east-west (x) direction and C be the variance in the north-south (y) direction. The covariance is B. Then the covariance matrix is as follows:

$$\Sigma=\begin{pmatrix} A & B \\ B & C \end{pmatrix}$$

Given this matrix the lengths of the semi-major and semi-minor axes are given by:

$$\text{Length of Semi-Major Axis} = \sqrt{5.991 * \left(A+C+\sqrt{(A-C)^2+4B^2} \right) /(2n)}$$

$$\text{Length of Semi-Minor Axis} = \sqrt{5.991 * \left(A+C-\sqrt{(A-C)^2+4B^2} \right) /(2n)}$$

The angle ψ (measured clockwise from North) of the Semi-Major Axis is

$$\theta = 90° - [\text{Arcsin}(2B/\sqrt{(A-C)^2+4B^2})]/2 \qquad \text{if } A \geq C$$

$$\text{or} \qquad \theta = [\text{Arcsin}(2B/\sqrt{(A-C)^2+4B^2})]/2 \qquad \text{if } A<C$$

Applying this to the matrices referenced in this report:

For $\Sigma = \begin{pmatrix} 4.0 & 0.3 \\ 0.3 & 0.3 \end{pmatrix}$ (used without reference to n, i.e. n=1)

We have: Angle to the Semi-Major Axis, θ=85.4°
Semi-Major Axis = 4.91
Semi-Minor Axis = 1.29

For $\Sigma = \begin{pmatrix} 0.2 & 0.3 \\ 0.3 & 2.0 \end{pmatrix}$ (used without reference to n, i.e. n=1)

We have: Angle to the Semi-Major Axis, θ=9.22°
Semi-Major Axis = 3.50
Semi-Minor Axis =.95

For $\Sigma = \begin{pmatrix} 2.00 & 0.15 \\ 0.15 & 0.15 \end{pmatrix}$ (used without reference to n, i.e. n=1)

We have: Angle to the Semi-Major Axis, θ=85.4°
Semi-Minor Axis = 3.47
Semi-Major Axis = .91

For $\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

We have: Angle to the Semi-Major Axis, NONE - it is circular
Lengths of axes depend on sample size, n, as follows

Sample Size	Semi-Major Axis	Semi-Minor Axis
5	1.09	1.09
10	.77	.77
20	.55	.55
50	.37	.37
100	.24	.24

For $\Sigma = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}$

We have: Angle to the Semi-Major Axis, θ=67.5
Lengths of axes depend on sample size, n, as follows

Sample Size	Semi-Major Axis	Semi-Minor Axis
5	2.64	.45
10	1.87	.32
20	1.32	.23
50	.83	.14
100	.59	.10

For $\Sigma = \begin{pmatrix} 1.0 & 0.7 \\ 0.7 & 1.0 \end{pmatrix}$

We have: Angle to the Semi-Major Axis, θ=45°
Lengths of the axes depend on sample size, n, as follows

Sample Size	Semi-Major Axis	Semi-Minor Axis
5	1.43	.60
10	1.01	.42
20	.71	.30
50	.45	.19
100	.32	.13

END

DTIC

5-86